

Mathematical Analysis of System Dynamics Model of a Peer-to-Peer Network

Vincent J. Papia

April 26, 2006

Abstract

This project provides a mathematical analysis of a seven-component ODE model of a peer-to-peer file sharing network. A steady state analysis was performed on the system. Under certain conditions, the steady state was shown to be unique. In addition, numerical simulations were performed using Simulink, a MATLAB package. Steady solutions and periodic solutions were found.

Contents

1	Introduction and Summary	2
2	Background	2
2.1	Peer Network Attractiveness	3
2.2	Network Attractiveness Equations	6
2.2.1	Peer Network Attractiveness Portrayed by the Media	6
2.2.2	Network Attractiveness as Perceived by Nonusers	6
2.2.3	Peer Network Attractiveness as Perceived by Users	7
2.3	Network Users	7
2.4	Shared Content	8
2.5	Total Shared Bandwidth	10
2.6	Average Traffic Per Successful Download	10
3	Analysis	11
3.1	Steady State Analysis	11
3.2	System Simplification	13
3.3	Steady State Existence and Uniqueness	14
4	Results	17
4.1	Simulink	17
4.2	Simulink Setup	18
4.3	Simulink Results	19
4.3.1	Solution Curves	20
A	Lookup Function Vectors	33
B	Simulink Diagram	34

1 Introduction and Summary

Peer-to-peer networking and file-sharing are phenomena that have reached almost universal popularity in recent history. The paper *A Resource Based Analysis of Peer-to-Peer Technology* [2] by Oleg Pavlov describes a system dynamics model of a peer-to-peer network. This model provides the system of seven, coupled, nonlinear, ordinary differential equations that is described and analyzed in this paper. A steady state analysis was performed on the system, yielding results that led to a steady state solution for the system. Under additional assumptions, the steady state solution was found to be unique. Numerical simulations of the model were performed using Simulink, a MATLAB package.

Two sets of results were obtained from Simulink. One set of results was produced by using the parameters and assumptions from the system dynamics model. Using these parameters, a steady state solution was not found. Instead, periodic behavior was found. Using parameters different from the ones given in the system dynamics model, a steady state solution to the system was found.

2 Background

In this section, the model of a peer-to-peer network will be discussed, as well as the formulation of a set of differential equations that model the behavior of a peer-to-peer network. A peer-to-peer network is a network of computers that are connected via a client program. Each computer that connects to this network acts as a separate node, contributing its own content (media files) and bandwidth (the amount of information that can pass across a line of communication in a given amount of time) to the overall network. Peer-to-peer networks are typically used for the sharing of media files such as MP3 files (music files), and AVI files (video files). Users of a particular peer-to-peer network download the client, usually hosted on a web site maintained by the developers of the client, install the client to their computer's hard drive, and then connect to the network using the client. Upon connection to the network, users can choose a user ID along with which files they own that they wish to share. Users can then search for files other users are sharing using the client's search engine. When a user finds a file he or she wishes to download, a download request will be initiated by whatever means the particular client he or she is using employs. Once a download request is made, a connection will be attempted between the destination and source of the file. After the connection has been made, the file will be transferred bit by bit from the source to the destination. This is the essence of a peer-to-peer network.

The ODE model we will develop is based on the model designed by Professors Oleg V. Pavlov and Khalid Saeed described in the papers *A Resource Based Analysis of Peer-to-Peer Technology* [2] and *Dynamic Analysis of an Institutional Conflict: Copyright Owners against Online File Sharing* [1]. This model has seven components. Three of these describe the attractiveness of the network as perceived by users of the network, nonusers, and the media. The remaining four components model behavior of Network Users, Shared Content, Total Shared Bandwidth, and Average Traffic per Successful Download. There are a number of variables and constants in the equations that model

the behavior of the seven sectors. Some variables and constants are exclusive to a certain sector while a few appear in each equation and are very influential in the overall behavior of the system. The values for each constant are given in a table in the Results section (section 4).

The seven state variables that represent the seven components of the model are:

State Variable	Description
a_m	The attractiveness of the network as perceived by the media
r_n	The attractiveness of the network as perceived by nonusers
r_u	The attractiveness of the network as perceived by users
U	The number of computers with unique user IDs connected to the network
C	The total shared content on the network
B	The total shared bandwidth of the network
t^{av}	The average traffic per successful download

2.1 Peer Network Attractiveness

Before discussion of the formulation of the first three equations, it will be necessary to describe the peer network attractiveness variable a since it appears in all three of them. Important characteristics of an attractive network include a content pool of interesting files, quick downloads, and high download success rates. If a network is lacking in any of these three areas, its attractiveness will drop accordingly. Because of this relationship, a is determined by forming the product of three variables that represent the three characteristics of an attractive network: content attractiveness a_c , latency acceptance a_l , and the probability of a successful download p_{sd} . Content attractiveness and the probability of a successful download are straightforward. Latency acceptance is related to the quickness of downloads. Latency is the amount of time it takes for a packet of information to travel from destination to source over a communication line. It is essentially a measure of the responsiveness of a network. Latency acceptance measures a user's level of patience with the latency of a network. So, peer network attractiveness is given by:

$$a = a_c \cdot a_l \cdot p_{sd} \quad (1)$$

These three quantities are each determined by a separate function, each of which will be described below:

$$\begin{aligned} a_c &= f_c(s) \\ a_l &= f_l(l) \\ p_{sd} &= 1 - p_{ud}(\lambda) \end{aligned}$$

The input to $f_c(s)$ is the relative richness of the network s . The relative richness of the network is related to the content available on the network and is given by the ratio of content of interest reachable by a user c_r to the typical user collection c_p , so $\frac{c_r}{c_p}$. The

typical collection c_p is assumed to be a constant, and content of interest reachable by a user is a variable given by:

$$c_r = \alpha_c \cdot n_v \cdot s^{av} \quad (2)$$

For this equation, the number of reachable nodes (user computers) n_v , is the largest number of computers any individual computer can search for files from and download from and is assumed to be constant. The value α_c is the share of interesting content. This value accounts for the dynamic nature of node connections by scaling the constant n_v by some value between 0 and 1. For this model, α_c is assumed constant. The variable s^{av} measures average shared content and is given by a ratio of shared content C , to network users U , so C/U . So s is given by:

$$s = \frac{\alpha_c n_v C}{c_p U} \quad (3)$$

Network users and shared content are detailed in sections 2.3 and 2.4 on pages 7 and 8 respectively. The properties of $f_c(s)$ are:

$$0 \leq f_c(s) \leq 1$$

$$f_c(0) = 0$$

$$f'_c(s) > 0$$

$$f''_c(s) < 0$$

$$\lim_{s \rightarrow \infty} f_c(s) = 1$$

Since content attractiveness is a function of the relative richness of the network, it is sensible for its value to be zero when there is no content of interest ($c_r = 0$) available on the network, i.e. $s = 0$. Also, content attractiveness should be increasing as s increases since as the richness of the network increases, so should its attractiveness. It is also sensible for content attractiveness to increase quickly at first and then slowly since when the richness of the network first begins its increase from zero, there should be a large increase of attractiveness. After s increases to the point where it is rare to introduce new and interesting content to the network, the increase will lose its intensity.

The input to latency acceptance, $f_l(l)$ is latency. Latency, l , is defined as the amount of time it takes for a packet of information to cross a communication path from sender to receiver. Latency l is determined by the equation:

$$l = (1 + l_{df}(\lambda))l_n \quad (4)$$

The function $l_{df}(\lambda)$ is the delay factor and is dependent on λ , the network utilization variable. This function is the principal component in determining latency. It measures the relationship between network utilization and latency. The variable λ measures the utilization of the network. It is the ratio of traffic requests to bandwidth, or T/B . The differential equation for total shared bandwidth is given by (28). Traffic requests, T is given by:

$$T = d_u U t^{av} \quad (5)$$

Where U is the variable for network users and t^{av} is average traffic per successful download, which will be described in section 2.6 on page 10. The value d_u represents download requests per user per day and is a constant. So the full expression for λ is:

$$\lambda = d_u t^{av} \frac{U}{B} \tag{6}$$

Returning to the function $l_{df}(\lambda)$, its properties are:

$$l_{df}(0) = 0$$

$$l'_{df}(\lambda) > 0$$

$$l''_{df}(\lambda) > 0$$

When network utilization is zero, there are either no download requests, no users, no traffic, or any combination of the three. It is sensible for latency to be zero when utilization is zero because there is essentially no activity on the network. It is also sensible for latency to increase quickly as network utilization increases, since increased network utilization implies increased network activity, and thus more of a strain on the network's resources. The constant l_n is the normal latency, which is a benchmark value for latency of the network.

Returning to the latency acceptance function, its properties are:

$$0 \leq f_l(l) \leq 1$$

$$f_l(1) = 1$$

$$f'_l(l) < 0$$

$$f''_l(l) > 0$$

As latency increases, the network's responsiveness will become lower so user tolerance of latency should decrease. It also makes sense that this tolerance will decrease rapidly as latency gets too high.

The final function that helps determine peer network attractiveness is the probability of an unsuccessful download. This function's input is λ . Its properties are:

$$0 \leq p_{ud} \leq 1$$

$$p_{ud}(0) = 0$$

$$p_{ud}(1) = 1$$

$$p'_{ud}(\lambda) > 0$$

$$\lim_{\lambda \rightarrow 0} p'_{ud}(\lambda) = 0$$

$$\lim_{\lambda \rightarrow 1} p'_{ud}(\lambda) = 0$$

Since this function measures probability, these properties are logical. As network utilization approaches 1, the probability that there will be an unsuccessful download should

reach one and remain there, because the traffic requests will approach the amount of bandwidth available on the network. When this situation occurs, there is too much activity on the network and not enough bandwidth to handle it, thus increasing the probability of dropped downloads.

So the full equation for peer network attractiveness a in terms of our state variables is given by:

$$a = f_c \left(\frac{\alpha_c n_v C}{c_p U} \right) f_l \left(\left(1 + l_{df} \left(d_u t^{av} \frac{U}{B} \right) \right) l_n \right) \left(1 - p_{ud} \left(d_u t^{av} \frac{U}{B} \right) \right) \quad (7)$$

Using (6), this can be written as follows:

$$a = f_c \left(\frac{\alpha_c n_v C}{c_p U} \right) f_l \left((1 + l_{df}(\lambda)) l_n \right) (1 - p_{ud}(\lambda)) \quad (8)$$

2.2 Network Attractiveness Equations

The attractiveness of the network is modeled by the equations for Network Attractiveness as perceived by the media, Network Attractiveness as perceived by nonusers, and Network Attractiveness as perceived by users. These three equations affect three of the four other sectors of the Peer-to-Peer Network model.

2.2.1 Peer Network Attractiveness Portrayed by the Media

A powerful force that affects trends and popular thought is the media. Positive media coverage of a network can increase its attractiveness and thus bring new users, while negative media portrayal can cause a decrease in new user arrivals. The variable a_m represents peer network attractiveness portrayed by the media. The equation modeling the behavior of a_m over time is:

$$\frac{da_m}{dt} = \frac{a - a_m}{\tau_m} \quad (9)$$

Where a is the peer network attractiveness variable, and τ_m is the opinion formation delay by the media, or how long it takes for the media to form an opinion; it is a constant.

This differential equation modeling the behavior of the attractiveness as perceived by the media is a very simple one. It is a first order linear equation that models the decay of a_m to the peer network attractiveness variable a . It is difficult to assume that the attractiveness perceived by the media will always approach the peer network attractiveness. The dynamic nature of the media is not fully captured by this equation, and as such, any results obtained from this model are not as accurate in a realistic sense as they would be if this equation was more complex.

2.2.2 Network Attractiveness as Perceived by Nonusers

Network attractiveness as perceived by nonusers r_n , is an important deciding factor that usually determines whether a person will connect to a network or not. This equation models how attractive the network is to people who have never used it before. Not

surprisingly, the fractional media effectiveness quantity influences the behavior of this equation. This should be clear since one of the ways a person learns about peer-to-peer networks is through the media. The equation modeling r_n is:

$$\frac{dr_n}{dt} = (E_m a_m) \frac{a - r_n}{\tau_n} \quad (10)$$

Where a is the peer network attractiveness variable, τ_n is a constant quantifying the opinion formation delay by nonusers which is the amount of time it takes for a nonuser to form an opinion, E_m is a constant that quantifies the effectiveness of the media, and a_m is the peer network attractiveness portrayed by the media variable. Its differential equation is given by (9).

2.2.3 Peer Network Attractiveness as Perceived by Users

The final variable in the Network Attractiveness sector is peer network attractiveness as perceived by users, r_u . This variable represents the satisfaction of users with the particular network they are using. The equation that models the behavior of r_u is:

$$\frac{dr_u}{dt} = \frac{a - r_u}{\tau_u} \quad (11)$$

Where once again, a is the peer network attractiveness variable, and τ_u is a constant for the opinion formation delay by users, or how long it takes for a user to form an opinion.

From the model, the relationship between the three opinion formation delay constants is as follows:

$$\tau_m < \tau_u < \tau_n$$

This relationship is sensible since the media is quick to form an opinion and report on it. Users have first-hand experience with networks so their opinion will form quickly, but not as quickly as the media's. The longest delay is in the nonusers group. Nonusers rely on second-hand sources like the media and other users for their information on networks, thus they take longer in forming their opinion.

2.3 Network Users

A major factor in the behavior of a peer-to-peer network is the number of connected users. Too many users can exhaust the resources of the network, while too few can cause a decrease in the amount of available content since users typically bring content to the network upon connection. We define network users U to be the number of people with unique user IDs that are connected to the network at any given time. The rate at which U changes is the difference between the number of new users connecting to the network and the number of users leaving the network; it is given by:

$$\frac{dU}{dt} = \text{new on-line users} - \text{attrition} \quad (12)$$

New on-line users is a measure of how many new arrivals to the network there are. The expression for new on-line users is given by:

$$\text{new on-line users} = u_a^t r_n U \quad (13)$$

Here, u_a^t is a constant that represents the typical adoption rate, the average rate at which people join the network each day. The variable r_n is the network attractiveness as perceived by nonusers and U is the amount of network users. The quantity $u_a^t r_n$ is called the fractional adoption rate.

Attrition is a measure of user departures from the network and is given by the expression:

$$\text{attrition} = U u_d^t (1 + g_d(r_u)) \quad (14)$$

In this equation u_d^t is the typical departure rate, which is the average rate at which people leave per day. The function $g_d(r_u)$ measures the effect of peer network attractiveness as perceived by users on the user departure rate. Its properties are:

$$\begin{aligned} g_d(1) &= 1 \\ g_d(r_u) &> 0 \\ g_d'(r_u) &< 0 \\ g_d''(r_u) &> 0 \end{aligned}$$

As attractiveness perceived by users increases, it makes sense that the departure rate relative to this attractiveness should decrease. It is also sensible that it should decrease quickly at first and then more slowly, since the initial increase from zero attractiveness will likely cause users to appreciate the network much more than when its attractiveness was zero. As users become accustomed to the attractiveness, the departure rate will continue to decrease, but at a much slower rate. The quantity $u_d^t (1 + g_d(r_u))$ is called the adjusted natural attrition rate.

So the equation that models U is:

$$\frac{dU}{dt} = U [u_a^t r_n - u_d^t (1 + g_d(r_u))] \quad (15)$$

2.4 Shared Content

Like the users sector, shared content C is another important aspect to observe when modeling a peer-to-peer network. Content is a driving factor that, when plentiful, will drive up user interest and increase attractiveness of the network, since there is a greater pool of resources to download from. Conversely, a lack of content will drive down interest when users are constantly searching for files with no results. The behavior of C is modeled by:

$$\frac{dC}{dt} = \text{added content} - \text{withdrawn content} \quad (16)$$

Added content is the amount of downloadable content a new user brings to the network upon arrival and is given by the equation:

$$\text{added content} = c^{av} \cdot \text{new on-line users} \quad (17)$$

Where c^{av} is the average new user contribution and is given by:

$$c^{av} = c^{max} \cdot \kappa \quad (18)$$

Where c^{max} is a constant that represents the maximum new user file contribution. The variable κ is the contribution fraction and it signifies the average fractional content contribution of an individual user to the network. It is given by:

$$\kappa = 1 - f^{or} f(U) \quad (19)$$

Where f^{or} is a constant representing the original free-riding fraction and $f(U)$ is the free-riding multiplier, which accounts for users who connect to a network and remain only until they are done downloading the files they searched for, sharing no files of their own. The effect of free-riding on the network is measured by the free-riding multiplier $f(U)$, scaled by the original free-riding fraction f^{or} , a constant. The input to this function is not exactly U , but a value of U scaled by the constant $U^{initial}$ which is an idealized value for the amount of users originally connected to the network. The free-riding multiplier's properties are:

$$\begin{aligned} f(U) &\geq 1 \\ \lim_{U \rightarrow \infty} f(U) &= \frac{1}{f^{or}} \\ f(U^{initial}) &= 1 \\ f'(U) &> 0 \end{aligned}$$

As the number of users connected to the network increases, it is realistic that the free-riding multiplier should increase, since there are more users who could be free-riding. Also, as the number of users tends to infinity, it is logical that the contribution fraction κ should approach zero since the free-riding effect increases as the number of users increases.

Withdrawn content is a measure of the content that is removed from the content pool when users disconnect from the network. It is given by:

$$\text{withdrawn content} = s^{av} \cdot \text{attrition} \quad (20)$$

The variable s^{av} is the average shared content of the network, or content per user. It is given by the ratio of content to users, so C/U . Attrition, which was described in the Users section is given by (14).

The equation that models shared content then is:

$$\frac{dC}{dt} = c^{max}(1 - f^{or} f(U))u_a^t r_n U - \frac{C}{U}u_d^t U(1 + g_d(r_u)) \quad (21)$$

Which simplifies to:

$$\frac{dC}{dt} = c^{max}(1 - f^{or} f(U))u_a^t r_n U - u_d^t C(1 + g_d(r_u)) \quad (22)$$

2.5 Total Shared Bandwidth

Bandwidth is defined as the amount of information that can pass through a communication line connected to a network through a modem or other device. In this model, we chose B to represent the total shared bandwidth of all users connected to the network. The equation modeling the behavior of B over time is similar to the equation for C and is given by:

$$\frac{dB}{dt} = \text{bandwidth contribution} - \text{bandwidth loss} \quad (23)$$

Bandwidth contribution is the amount of bandwidth each new user brings when he or she connects to the network and is equal to:

$$\text{bandwidth contribution} = \text{new on-line users} \cdot b_{avg} \quad (24)$$

Where b_{avg} is the average new node (user's computer) contribution of bandwidth which is given by:

$$b_{avg} = b_t \cdot \kappa \quad (25)$$

Here, b_t is a constant representing typical node bandwidth and κ is the contribution fraction and is given by (19). As before, new on-line users is given by (13).

Bandwidth loss is the bandwidth that is lost when a user disconnects from the network. It is given by:

$$\text{bandwidth loss} = b^{av} \cdot \text{attrition} \quad (26)$$

The variable b^{av} is the average shared bandwidth and is the ratio of shared bandwidth to users, or B/U . Again, attrition is given by (14).

So, the equation that models the total shared bandwidth of the network is:

$$\frac{dB}{dt} = b_t(1 - f^{or} f(U))u_a^t r_n U - \frac{B}{U}u_d^t U(1 + g_d(r_u)) \quad (27)$$

Which simplifies to:

$$\frac{dB}{dt} = b_t(1 - f^{or} f(U))u_a^t r_n U - u_d^t B(1 + g_d(r_u)) \quad (28)$$

2.6 Average Traffic Per Successful Download

The last sector in the peer-to-peer network model is the average traffic per successful download, t^{av} . Traffic is the passage of data through a network. The measure of traffic is essentially a measure of all activity occurring throughout the network, whether by searching for files, making download requests, or downloading files. The equation modeling the change in average traffic per successful download is:

$$\frac{dt^{av}}{dt} = \text{traffic adjustment} = t^* - t^{av} \quad (29)$$

The quantity t^* is traffic per download adjusted for traffic congestion. Traffic congestion is a situation where there is a large amount of network activity that the network cannot

adequately handle. It is a non-linear function dependent on network utilization λ and traffic elasticity α_{te} . It is determined by:

$$t^* = (1 + p_{ud}(\lambda))^{\alpha_{te}} \cdot (1 + l_{df}(\lambda))^{\alpha_{te}} \cdot t_b \quad (30)$$

This equation involves the two functions $p_{ud}(\lambda)$ and $l_{df}(\lambda)$, which were described in section 2.1. The parameter α_{te} is traffic elasticity, which is a measure of the sensitivity of traffic volume to congestion in the network. It was chosen to be constant for this model, though it can vary between 0 and 1. Finally, the constant t_b is base traffic per download. So the full equation for average traffic per successful download is:

$$\frac{dt^{av}}{dt} = (1 + p_{ud}(\lambda))^{\alpha_{te}} \cdot (1 + l_{df}(\lambda))^{\alpha_{te}} \cdot t_b - t^{av} \quad (31)$$

Recall from (6) that λ depends on U , t^{av} , and B .

3 Analysis

The analysis of this system began with a steady state analysis. After finding the steady state equations, all seven equations were reduced to one equation in terms of λ , the network utilization variable. The seven differential equations are:

$$\frac{da_m}{dt} = \frac{a - a_m}{\tau_m} \quad (32)$$

$$\frac{dr_n}{dt} = \frac{a - r_n}{\tau_n} (E_m a_m) \quad (33)$$

$$\frac{dr_u}{dt} = \frac{a - r_u}{\tau_u} \quad (34)$$

$$\frac{dU}{dt} = U[u_a^t r_n - u_d^t (1 + g_d(r_u))] \quad (35)$$

$$\frac{dC}{dt} = c^{max} (1 - f^{or} f(U)) u_a^t r_n U - u_d^t C (1 + g_d(r_u)) \quad (36)$$

$$\frac{dB}{dt} = b_t (1 - f^{or} f(U)) u_a^t r_n U - u_d^t B (1 + g_d(r_u)) \quad (37)$$

$$\frac{dt^{av}}{dt} = (1 + p_{ud}(\lambda))^{\alpha_{te}} \cdot (1 + l_{df}(\lambda))^{\alpha_{te}} \cdot t_b - t^{av} \quad (38)$$

3.1 Steady State Analysis

After gaining a basic understanding of each of the equations, a steady state analysis was performed. In order for a steady state to occur in the first three equations, it is necessary that each variable be equal to a , the peer network attractiveness variable. So

$$r_u = r_n = a_m = a$$

From the network users equation, two conditions for the steady state were found. One condition is when $U = 0$, and there is nobody connected to the network. This steady

state is not very interesting and describes a dead network; so we have ruled this condition out. The other condition is when the fractional adoption rate is equal to the adjusted natural attrition rate:

$$u_a^t r_n = u_d^t (1 + g_d(r_u)) \quad (39)$$

Also, r_u and r_n are equal to a , so by substitution:

$$u_a^t a = u_d^t (1 + g_d(a))$$

This relationship implies that the rate at which people are connecting to the network is equal to the rate at which people are disconnecting from the network. In other words, this means that anytime somebody disconnects from the network, someone else will connect to the network at the exact same time, keeping U constant.

In the content sector, it was found that shared content will be constant either when $U = 0$, which implies $C = 0$, signifying that there are no users connected to the network and since users bring content with them upon connection to the network, no shared content on the network, which again is not allowed. Otherwise, there is a steady state when $U \neq 0$ and:

$$c^{max} (1 - f^{or} f(U)) u_a^t r_n U = u_d^t C (1 + g_d(r_u))$$

By using the relationship from (39), this equation simplifies to:

$$C = U c^{max} (1 - f^{or} f(U)) = U c^{max} \kappa$$

This value for C implies that in order for the system to reach a steady state, the total number of files shared is the maximum new user contribution scaled by the contribution fraction.

Since the total shared bandwidth equation is so similar to the equation modeling total shared content, the steady state conditions are similar. A steady state occurs either when $U = 0$ and thus $B = 0$, since users contribute their bandwidth to the network upon connection, which is again not allowed, or when $U \neq 0$ and:

$$b_t (1 - f^{or} f(U)) u_a^t r_n U = u_d^t B (1 + g_d(r_u))$$

This equation can also be simplified by using the relationship from (39)

$$B = U b_t (1 - f^{or} f(U)) = U b_t \kappa$$

The steady state value for B implies that the average bandwidth is the typical node bandwidth b_t scaled by the contribution fraction.

The equation modeling average traffic per successful download is the final equation. The only steady state condition occurs when:

$$t^{av} = (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} t_b$$

The value for t^{av} at its steady state implies that the average traffic per download is the base traffic per download scaled by fractional powers of one plus the probability of an unsuccessful download and one plus the delay factor.

Collecting the results for $U \neq 0$, we have:

$$a = f_c \left(\frac{\alpha_c n_v C}{c_p \bar{U}} \right) f_l ((1 + l_{df}(\lambda)) l_n) (1 - p_{ud}(\lambda)) \quad (40)$$

$$r_u = r_n = a_m = a \quad (41)$$

$$u_a^t a = u_d^t (1 + g_d(a)) \quad (42)$$

$$C = U c^{max} \kappa \quad (43)$$

$$B = U b_t \kappa \quad (44)$$

$$t^{av} = (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} t_b \quad (45)$$

3.2 System Simplification

Using the steady state equations, a further simplification of the system is possible. It turns out that the steady state conditions can be reduced to a single equation in λ . A value for a can be determined from (42). It will be shown in the beginning of section 3.3 that (42) determines a unique value of a . Given this value of a , we will be able to write (40) so that only λ appears on the right hand side.

First, by using the steady state condition on C (eq. (43)), eq. (40) in terms of κ and λ is:

$$a = f_c \left(\frac{\alpha_c n_v c^{max} \kappa}{c_p} \right) f_l ((1 + l_{df}(\lambda)) l_n) (1 - p_{ud}(\lambda)) \quad (46)$$

Recall that λ is given by $\lambda = d_u t^{av} \frac{U}{B}$. By using (44):

$$\lambda = \frac{d_u t^{av}}{b_t \kappa}$$

is obtained. By taking the steady state condition for t^{av} , (45), and multiplying both sides by $\frac{d_u}{b_t \kappa}$:

$$\frac{d_u t^{av}}{b_t \kappa} = (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t \kappa}$$

is obtained. Recognizing that the expression on the left-hand side is λ , we obtain the following, which now involves only λ and κ :

$$\lambda = (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t \kappa}$$

This equation can be rearranged to give an expression for κ . This is done by multiplying both sides by κ and dividing both sides by λ :

$$\kappa = (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t \lambda}$$

This expression is in terms of only λ so κ can now be thought of as a function $\kappa = \Lambda(\lambda)$. So

$$\kappa = \Lambda(\lambda) \equiv (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t \lambda} \quad (47)$$

By substituting $\Lambda(\lambda)$ for κ in (46), a new expression where only λ appears on the right-hand side is gained. This expression is:

$$a = f_c \left(\frac{\alpha_c n_v c^{max} \Lambda(\lambda)}{c_p} \right) f_l((1 + l_{df}(\lambda))l_n)(1 - p_{ud}(\lambda))$$

All terms other than λ in this equation are constants. Consider this equation to be a function called $H(\lambda)$, so we have the following equation involving only a and λ :

$$a = H(\lambda) \equiv f_c \left(\frac{\alpha_c n_v c^{max} \Lambda(\lambda)}{c_p} \right) f_l((1 + l_{df}(\lambda))l_n)(1 - p_{ud}(\lambda)) \quad (48)$$

We will show that under our assumptions there is a unique value of a , determined from (42). Eq. (48) then is to be solved for λ . Below we show that under our assumptions, there is always at least one λ , $0 < \lambda < 1$, that satisfies (48). Furthermore, we show that under an additional assumption the value of λ is unique.

The following process provides values for each state variable. This process is summarized by:

1. Using the value for a obtained from (42) in (48), a unique value for λ can be determined.
2. Using (45), a value can be determined for t^{av} .
3. Also using the value for λ , a value for κ can be determined using (47).
4. From the definition of κ ($\kappa = 1 - f^{or} f(U)$) a value can be determined for U .
5. Using the value of U , values can be determined for both C and B from the steady state equations for each.
6. Finally, since $r_n = r_u = a_m = a$, a value for each of the network attractiveness variables can be determined.

3.3 Steady State Existence and Uniqueness

It is easy to show that under the assumption:

$$0 < \frac{u_d^t}{u_a^t} < \frac{1}{2}$$

there is a unique value of a , $0 < a < 1$, that satisfies (42). Simply rewrite (42) as:

$$a = \frac{u_d^t}{u_a^t} (1 + g_d(a))$$

and recall that the function $g_d(a)$ satisfies the conditions $g_d(a) > 0$, $g_d'(a) < 0$, and $g_d(1) = 1$. Using the assumption, we obtain the following:

$$\frac{u_d^t}{u_a^t} (1 + g_d(a)) > 0$$

and

$$\frac{u_d^t}{u_a^t}(1 + g_d(1)) = \frac{u_d^t}{u_a^t} \cdot 2 < 1$$

which imply that there is at least one value of a , $0 < a < 1$, that satisfies (42). Uniqueness follows from the fact that $g_d(a)$ is strictly decreasing.

The next question is whether or not a solution to $H(\lambda) = a$ exists for any given a , $0 < a < 1$. We will show that $H(0) = 1$ (more precisely, $\lim_{\lambda \rightarrow 0^+} H(\lambda) = 1$) and $H(1) = 0$. Continuity of H then gives us existence. Given that $l_{df}(0) = 0$, $f_l(1) = 1$, and $p_{ud}(0) = 0$, we have:

$$\lim_{\lambda \rightarrow 0^+} H(\lambda) = \lim_{\lambda \rightarrow 0^+} f_c \left(\frac{\alpha n_v c^{max}}{c_p} \Lambda(\lambda) \right)$$

Since $\lim_{\lambda \rightarrow 0^+} \Lambda(\lambda) = \infty$, we have:

$$\lim_{\lambda \rightarrow 0^+} H(\lambda) = \lim_{s \rightarrow \infty} f_c(s) = 1$$

At $\lambda = 1$, there is no need for limits, and the condition $p_{ud}(1) = 1$ gives $H(1) = 0$.

The value of λ that satisfies $H(\lambda) = a$ must be unique if $H'(\lambda) < 0$ for $0 < \lambda < 1$. This derivative is given by:

$$\begin{aligned} H'(\lambda) = & c_1 f'_c(c_1 \Lambda(\lambda)) \Lambda'(\lambda) f_l(1 + l_{df}(\lambda))(1 - p_{ud}(\lambda)) \\ & + f_c(c_1 \Lambda(\lambda)) f'_l(1 + l_{df}(\lambda)) l'_{df}(\lambda) (1 - p_{ud}(\lambda)) \\ & - f_c(c_1 \Lambda(\lambda)) f_l(1 + l_{df}(\lambda)) p'_{ud}(\lambda) \end{aligned} \quad (49)$$

From the properties of each of these functions, it is known that:

$$\begin{aligned} f_c(s) & \geq 0 \\ f'_c(s) & > 0 \\ f_l(l) & \geq 0 \\ f'_l(l) & < 0 \\ l'_{df}(\lambda) & > 0 \\ p'_{ud}(\lambda) & > 0 \end{aligned}$$

Since p_{ud} is a probability, it takes values between zero and one. Also notice that $c_1 > 0$ since it is composed of four positive constants. Therefore, all signs of all terms excluding the sign of $\Lambda'(\lambda)$ is known. The second and third addends in $H'(\lambda)$ are negative. However, because the sign of $\Lambda'(\lambda)$ is not known, the first addend's sign is also not known. Therefore, the sign of $\Lambda'(\lambda)$ is very important. If it is negative, then $H'(\lambda)$ is negative as well since all three addends are negative, implying that $H(\lambda)$ is strictly decreasing, leading to a single value of λ for any value of a . So if $\Lambda'(\lambda)$ is negative, then the process summarized at the end of section 3.2 determines a single steady state for this system. If $\Lambda'(\lambda)$ is positive, this introduces the possibility that $H'(\lambda)$ may be zero, since two of the addends are positive and one is negative. This leads to a possible sign change for $H'(\lambda)$.

If $H(\lambda)$ is not strictly increasing or decreasing, then it is possible that a single value of a could lead to multiple values for λ , thus leading to multiple steady states for the system. Therefore, the uniqueness of the steady state solution depends on the sign of $\Lambda'(\lambda)$. The function $\Lambda(\lambda)$ is given by (47). So the derivative of $\Lambda(\lambda)$ is:

$$\begin{aligned}\Lambda'(\lambda) &= \alpha_{te}(1 + p_{ud}(\lambda))^{\alpha_{te}-1} p'_{ud}(\lambda)(1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t} \frac{1}{\lambda} \\ &\quad + \alpha_{te}(1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}-1} l'_{df}(\lambda) \frac{t_b d_u}{b_t} \frac{1}{\lambda} \\ &\quad - (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t} \frac{1}{\lambda^2}\end{aligned}\tag{50}$$

Which can be rewritten as:

$$\begin{aligned}\Lambda'(\lambda) &= \alpha_{te} \frac{p'_{ud}(\lambda)}{(1 + p_{ud}(\lambda))} (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t} \frac{1}{\lambda} \\ &\quad + \alpha_{te} \frac{l'_{df}(\lambda)}{(1 + l_{df}(\lambda))} (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t} \frac{1}{\lambda} \\ &\quad - (1 + p_{ud}(\lambda))^{\alpha_{te}} (1 + l_{df}(\lambda))^{\alpha_{te}} \frac{t_b d_u}{b_t} \frac{1}{\lambda^2}\end{aligned}\tag{51}$$

Notice now that each term in this derivative now contains $\Lambda(\lambda)$. So by substitution, (51) becomes:

$$\begin{aligned}\Lambda'(\lambda) &= \alpha_{te} \frac{p'_{ud}(\lambda)}{(1 + p_{ud}(\lambda))} \Lambda(\lambda) + \alpha_{te} \frac{l'_{df}(\lambda)}{(1 + l_{df}(\lambda))} \Lambda(\lambda) - \frac{1}{\lambda} \Lambda(\lambda) \\ &= \Lambda(\lambda) \left(-\frac{1}{\lambda} + \alpha_{te} \left(\frac{p'_{ud}(\lambda)}{(1 + p_{ud}(\lambda))} + \frac{l'_{df}(\lambda)}{(1 + l_{df}(\lambda))} \right) \right)\end{aligned}\tag{52}$$

From this final form of $\Lambda'(\lambda)$, a condition can be found that will cause $\Lambda'(\lambda)$ to be negative. This condition is:

$$\alpha_{te} \lambda \left(\frac{p'_{ud}(\lambda)}{(1 + p_{ud}(\lambda))} + \frac{l'_{df}(\lambda)}{(1 + l_{df}(\lambda))} \right) < 1\tag{53}$$

Whenever this condition holds, $\Lambda'(\lambda)$ will be negative, making $H'(\lambda)$ negative, thus implying a unique value of λ that satisfies (48) for any given value of a , $0 < a < 1$.

Given a value of λ , the next step is to compute κ from (47). A complication arises if this value given by (47) is greater than one, because there will then not be a value of U that satisfies (19), meaning that no steady state exists. This has not proven to be a problem for the parameter values used, but warrants further investigation.

4 Results

The process summarized in the end of section 3.2 gives steady state values for each of the state variables. There are a number of parameters that were described in the Background section that must be taken into consideration when determining the steady state values. The following table gives the nominal values used for each of the constants in this model.

Parameter	Description	Value
τ_n	opinion formation delay for nonusers	100 days
τ_u	opinion formation delay for users	25 days
τ_m	opinion formation delay for the media	1 day
E_m	media effectiveness	2.3 dimensionless
u_a^t	normal adoption rate	0.02 dimensionless
u_d^t	typical attrition rate	0.001 dimensionless
f^{or}	original free riding fraction	0.09 dimensionless
$U^{initial}$	initial network users	2000 Users
n_v	number of reachable nodes	381 nodes
c_p	typical collection	500 files
α_c	share of interesting content	0.0025 dimensionless
c^{max}	max new user contribution	3000 files
b_t	typical node bandwidth	5.184×10^9 bits per day
d_u	download requests per user per day	1.87 files
l_n	normal latency	1 dimensionless
α_{te}	traffic elasticity	0.192 dimensionless
t_b	base traffic per download	15000 bits

In addition to these parameters, each of the functions described in the Background section were defined by pairs of input and output vectors. These vectors were determined based on the behavior that each function should exhibit. The vectors defining these functions are given in Appendix A.

4.1 Simulink

In order to obtain a better understanding of the peer-to-peer network model, as well as graphical representations of the solution curves, a Simulink representation of the model was built. Simulink is a block library tool used in the simulating of dynamic systems [3]. It is a software package used with MATLAB, a technical computing environment. Simulink was designed to simulate dynamic systems that change over time, a type of system dealt with in control theory. Within Simulink, there are blocks that can be used to represent functions, variables, constants, and integrators. These blocks were arranged and connected in such a way that allowed the entire peer-to-peer network system to be modeled. Once the representation was setup, Simulink could solve the system numerically and plot the solutions of the system over time. Because of this feature, and because it was relatively easy to setup, Simulink was an excellent choice to simulate this system of

differential equations.

4.2 Simulink Setup

A diagram of the Simulink representation can be found in Appendix B. In order to use Simulink, MATLAB must be installed on the computer the user is using. On a computer running Windows XP, at MATLAB's startup, a command prompt will be presented to the user. To start the Simulink package, the user must type "simulink" into the command prompt. This will open the Simulink Library Browser. From this browser window, a new system representation can be built or an existing one can be opened. To start a new system representation, the user must either click on the white paper icon with his/her mouse pointer, or from the top menu, click on file, then new, and then model. This will bring up an environment to build a new block diagram. Blocks from the Simulink Library Browser can be clicked and dragged with the mouse pointer into the new model window. There are a myriad of blocks that each perform specific functions.

In the diagram, simple arithmetic operations are taken care of by sum, product, and division blocks. The sum and product blocks can be found in both the "Commonly Used Blocks" or "Math Operations" section. The division block is found in the "Math Operations" section. Each of these blocks has a user-defined number of inputs, and a single output. By double-clicking on any one of these blocks, the properties of the blocks can be changed. For example, double-clicking on the sum block will bring up a menu where the operations can be specified, as well as the number of inputs. For operations specification in the sum block, the user must first enter a "-" character followed by the operations he/she wishes to use. For addition the "+" character is used; for subtraction, the "-" character is used. Similarly, in the division box, multiplication is specified with the "*" character, while division is specified with the "/" character. In the product block, only the number of inputs can be changed.

For each of the parameters in the system, constant blocks were used. The constant block can be found either in the "Commonly Used Blocks" or "Sources" sections of the Simulink Library Browser. By double-clicking on these blocks, a window will open and allow the user to enter the value he/she wish to use for this parameter. There is a single output that can send the value of the constant to whatever input is necessary.

In order to handle the six functions, lookup table blocks were used. The lookup table block can be found in the "Lookup Tables" section of the Simulink Library Browser. To specify the input and output values for a function, the user must double-click on the lookup table block and then enter the input and output vectors that define the function they wish to use. This block also features extrapolation for values beyond the range of the specified vectors. A number of different extrapolation methods are available, but for this system, the "Use End Values" option was used to determine output values for input values greater than the largest value from the input vector.

Since this diagram represents a system of differential equations, an integrator block was used. The integrator block can be found in both the "Commonly Used Blocks" and "Continuous" sections of the Simulink Library Browser. This block outputs the result of the integral of the input. The upper limit of integration is user-definable while the lower limit starts at time equal to zero. The upper limit can be changed in the menu bar.

The default value is 10.0. By double-clicking on the integrator block, the user will be presented with a menu that allows him/her to specify the initial value of the particular variable they are integrating.

The displaying of the results was handled in two ways. To get an immediate view of the solution curves, mainly for testing and debugging, a scope block was used. The scope block can be found in the “Commonly Used Blocks” and “Sinks” sections of the Simulink Library Browser. Double-clicking on a scope block will produce a graph representing the input as it changes over time. In order to send results to the MATLAB workspace, the “to workspace” block was used. This block can be found in the “Sinks” section of the Simulink Library Browser. This block sends its input to the MATLAB workspace as a vector. By double-clicking on the “to workspace” block, the user can name the output whatever he/she desires as well as choose the type of variable to be sent. For this system’s representation, each value was sent as an “Array”. To use the results, the user must simply type the name of the variable he/she specified into the MATLAB command prompt. From here, the results can be treated just as any other variable can be in MATLAB.

4.3 Simulink Results

In order for there to be a steady state solution to the system in Simulink, modifications were made to the original parameters. These modifications are:

Parameter	Description	Value
τ_n	opinion formation delay for nonusers	15 days
τ_u	opinion formation delay for users	7 days
E_m	media effectiveness	3.6 dimensionless

The resulting numerical values for the state variables given by Simulink using the modified parameters are:

State Variable	Description	Steady State Value
a	Peer network attractiveness	0.23917
λ	Network utilization	1.6328×10^{-4}
t^{av}	Average traffic per successful download	1.5185×10^4
κ	Contribution fraction	0.03357
U	Number of computers with unique user IDs connected to the network	2.7905×10^4
C	Total shared content on the network	2.8084×10^6
B	Total shared bandwidth of the network	4.8529×10^{12}

Remember that for the steady state, the variables r_n , r_u , and a_m are all equal to a .

4.3.1 Solution Curves

Figures 1 through 12 represent the solution curves of each of the equations when the modified values of the parameters were used. As the figures show, there is a steady state solution for the system under these modified parameters.

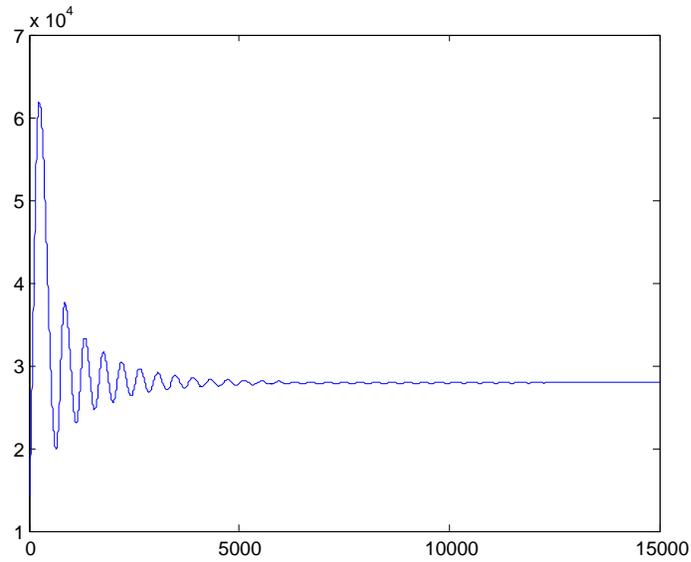


Figure 1: Network Users, U

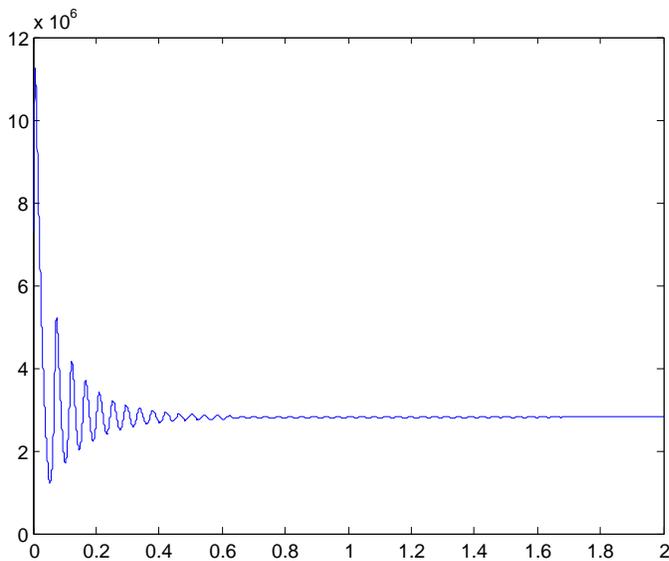


Figure 2: Total Shared Content, C

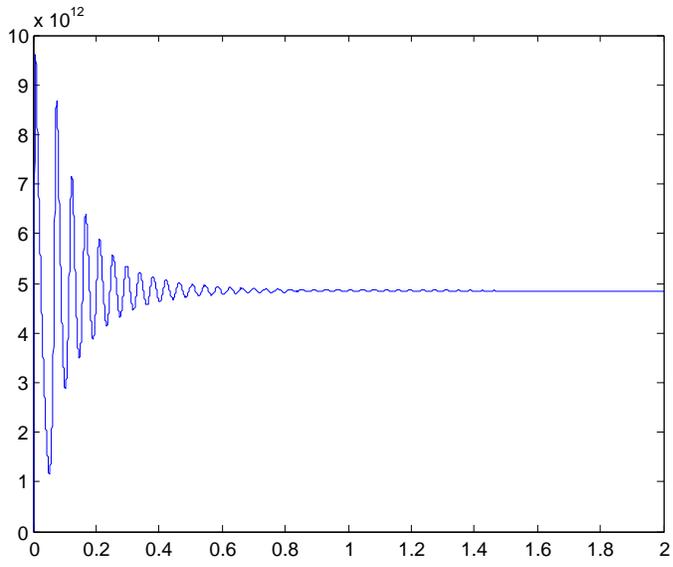


Figure 3: Total Shared Bandwidth, B

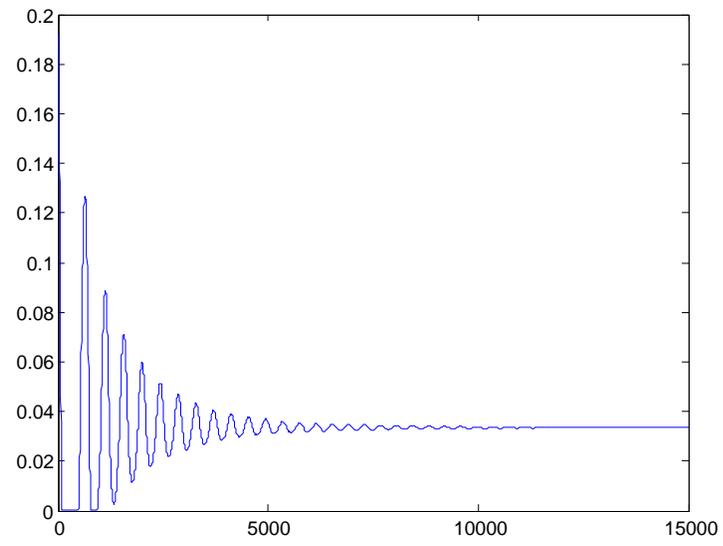


Figure 4: Contribution Fraction, κ

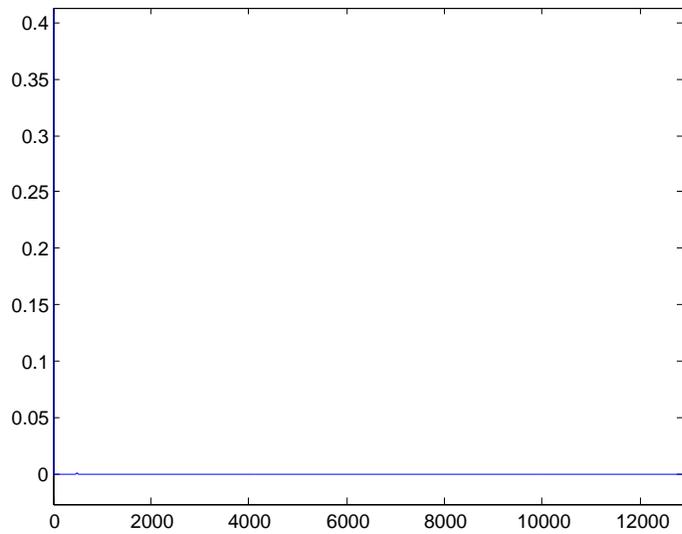


Figure 5: Network Utilization, λ

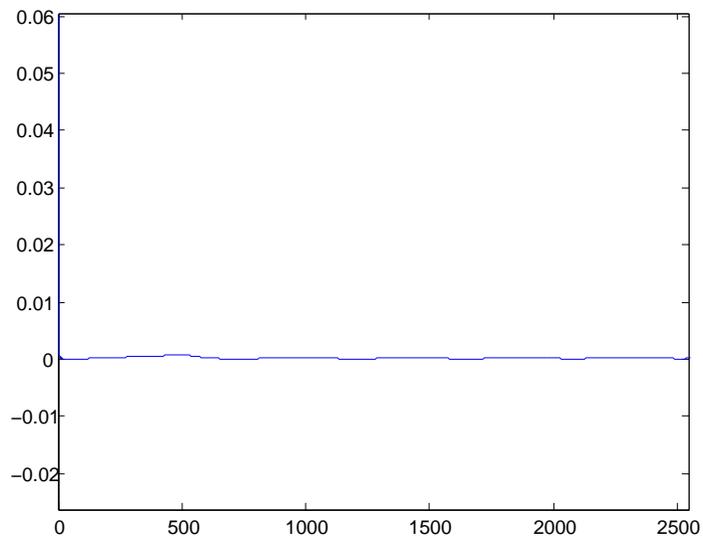


Figure 6: Network Utilization (Zoomed in), λ

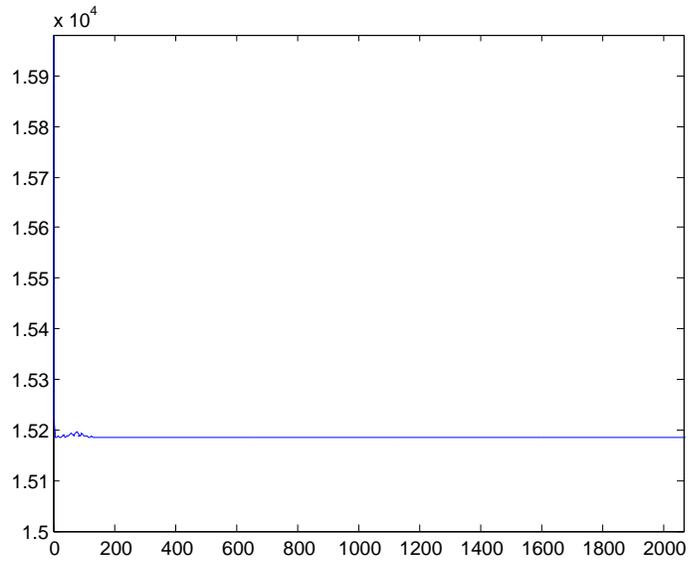


Figure 7: Average Traffic per Successful Download, t^{av}

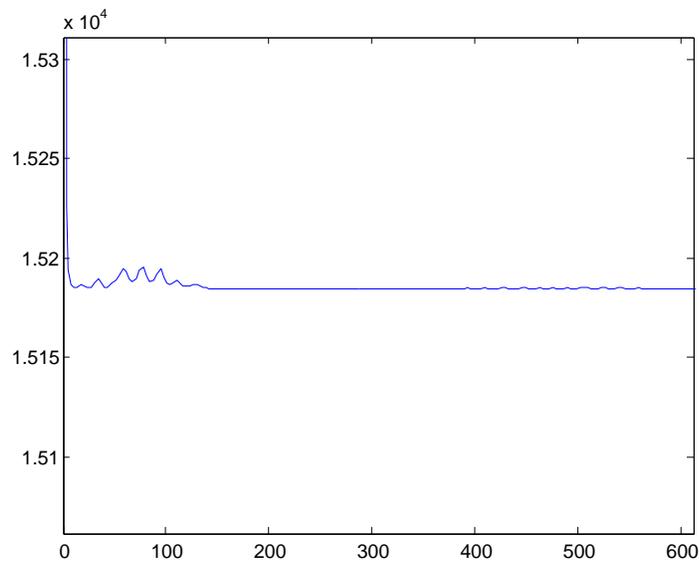


Figure 8: Average Traffic per Successful Download (Zoomed in), t^{av}

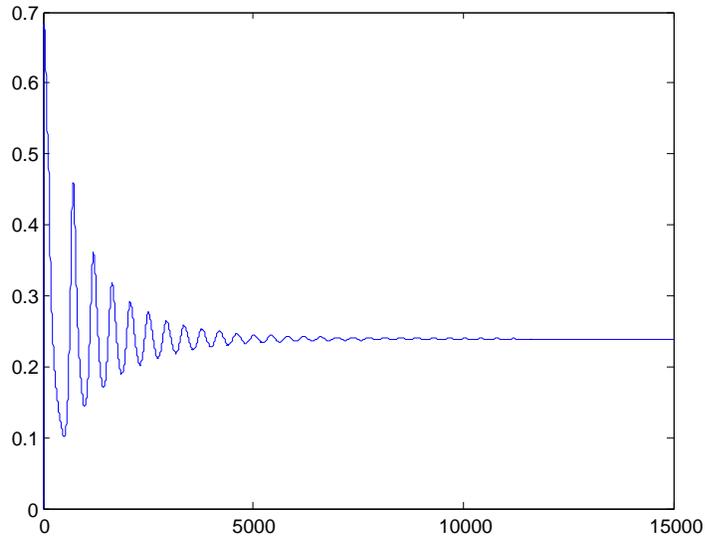


Figure 9: Peer Network Attractiveness, a

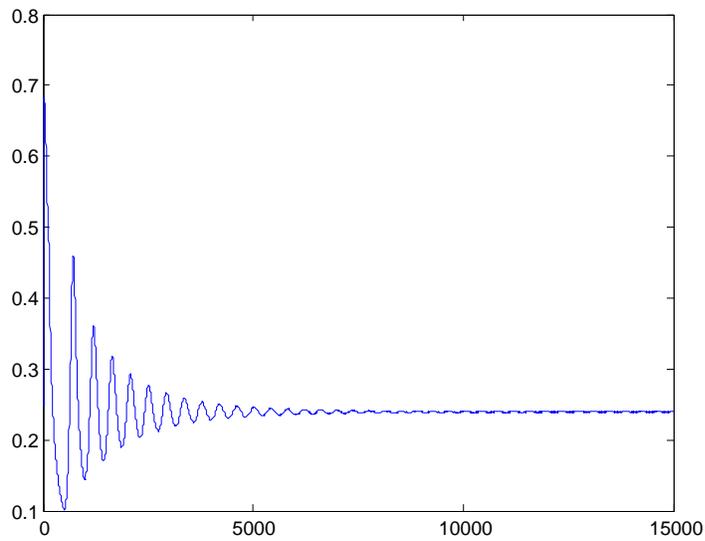


Figure 10: Network Attractiveness as Perceived by the Media, a_m

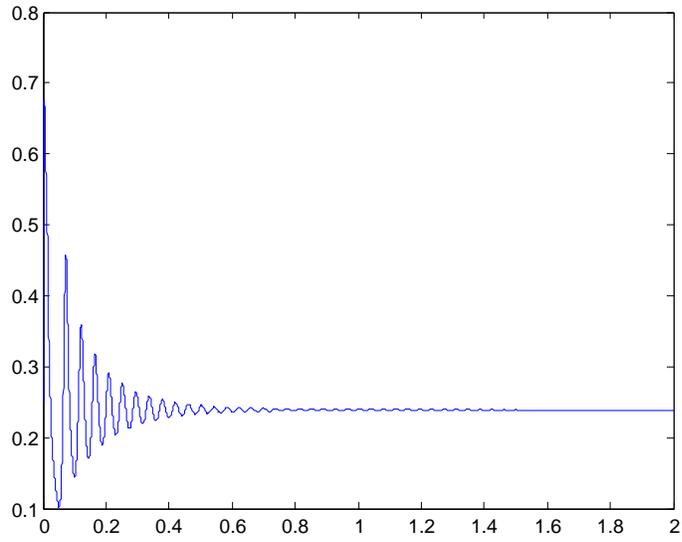


Figure 11: Network Attractiveness as Perceived by Users, r_u

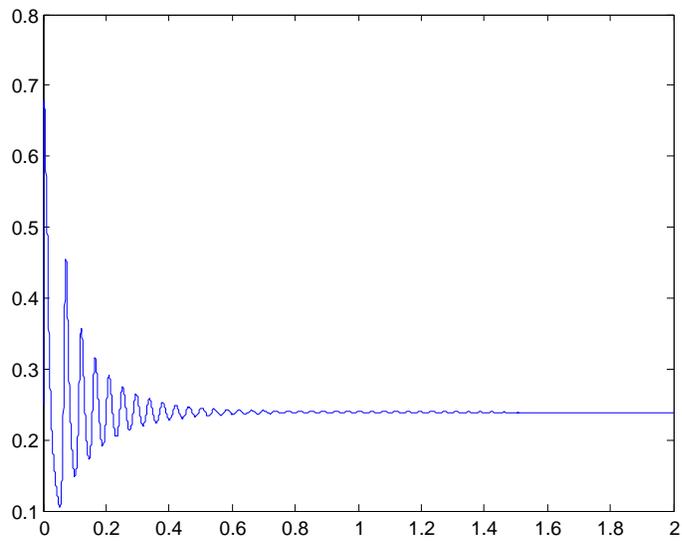


Figure 12: Network Attractiveness as Perceived by Nonusers, r_n

Figures 13 through 24 represent the solution curves of the state variables over time when the original parameters were used. Notice the periodic behavior in all solutions except for Average Traffic per Successful Download, which is nearly unchanged from the modified parameters.

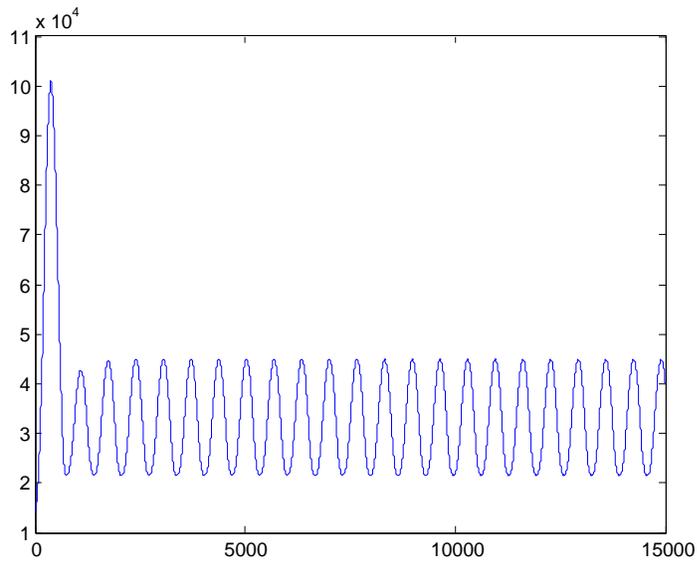


Figure 13: Network Users, U

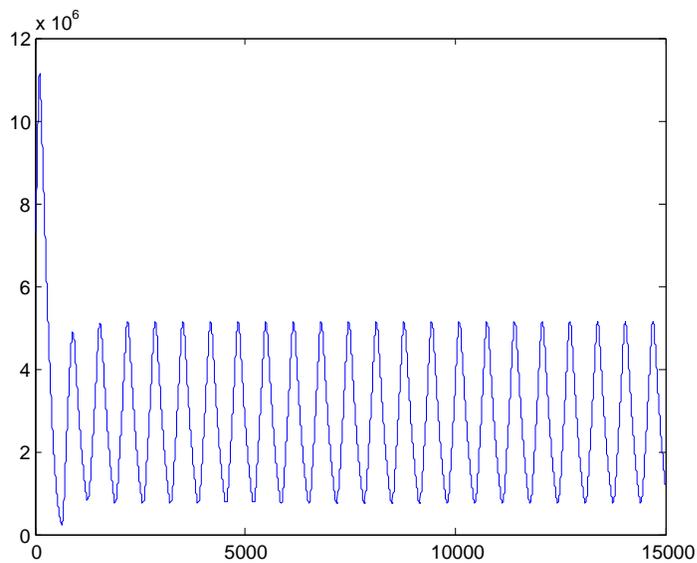


Figure 14: Total Shared Content, C

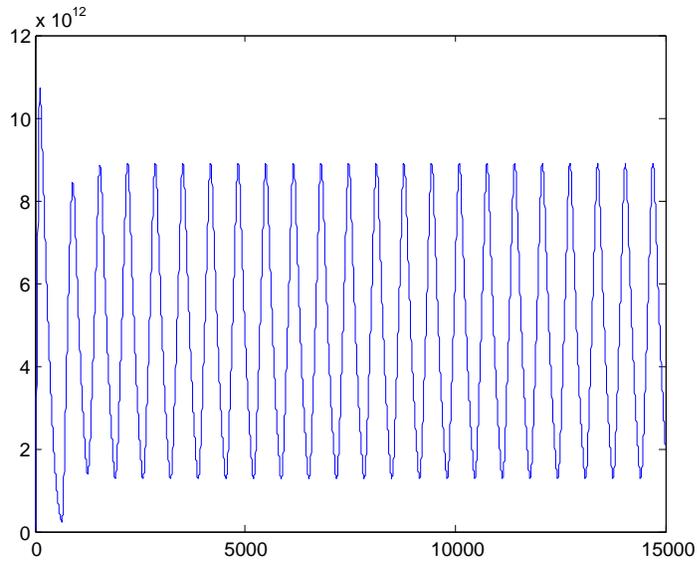


Figure 15: Total Shared Bandwidth, B

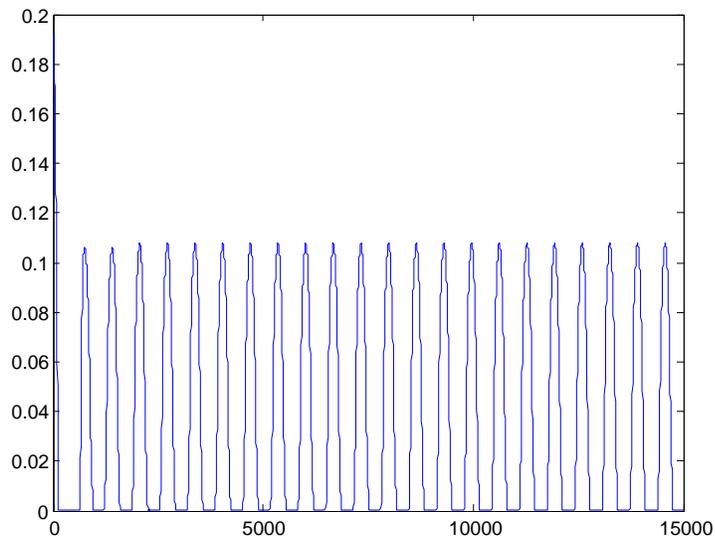


Figure 16: Contribution Fraction, κ

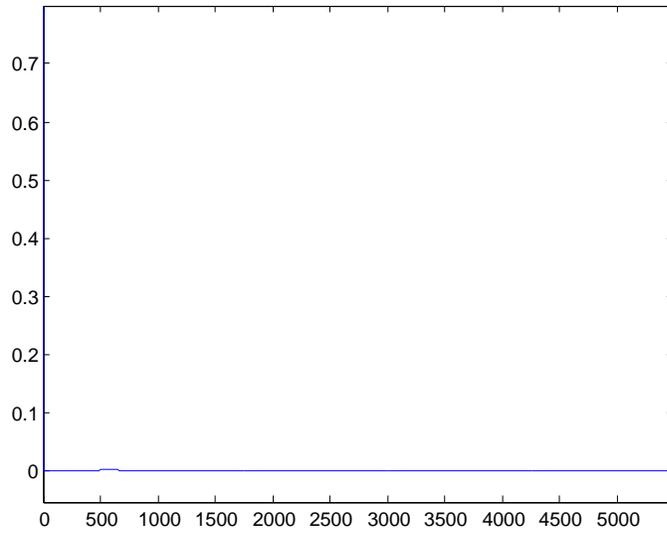


Figure 17: Network Utilization, λ

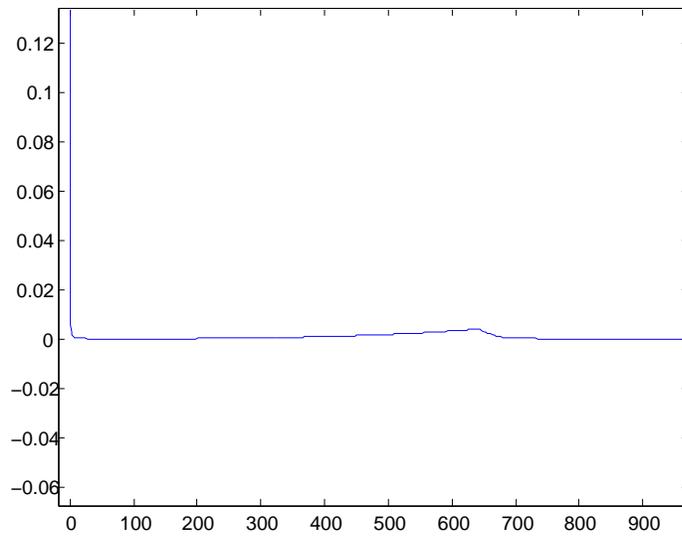


Figure 18: Network Utilization (Zoomed in), λ

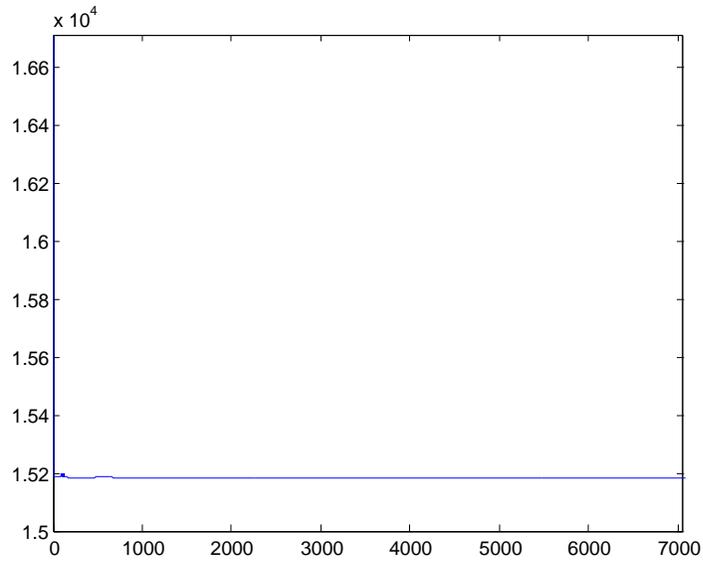


Figure 19: Average Traffic per Successful Download, t^{av}

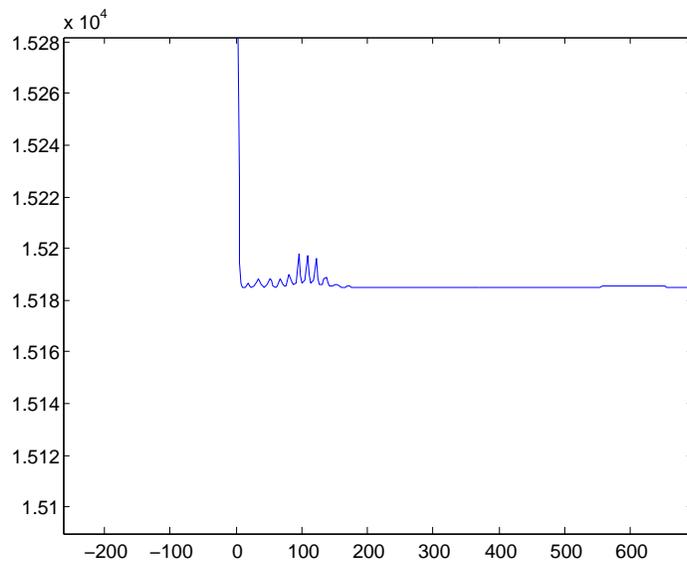


Figure 20: Average Traffic per Successful Download (Zoomed in), t^{av}

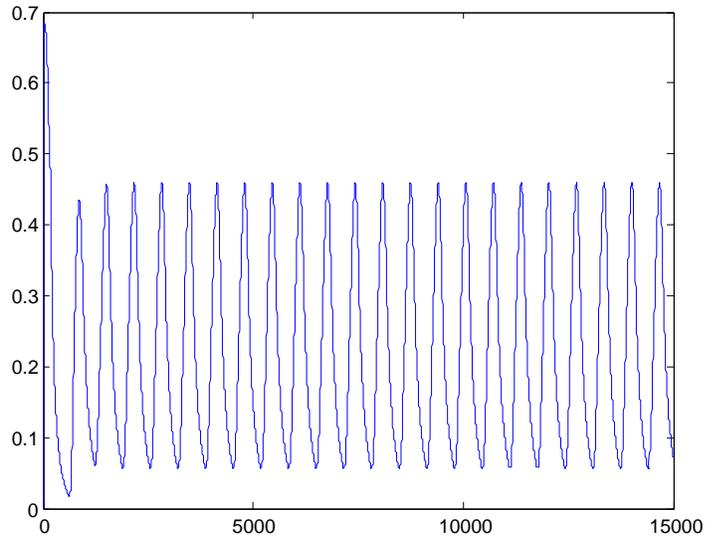


Figure 21: Peer Network Attractiveness, a

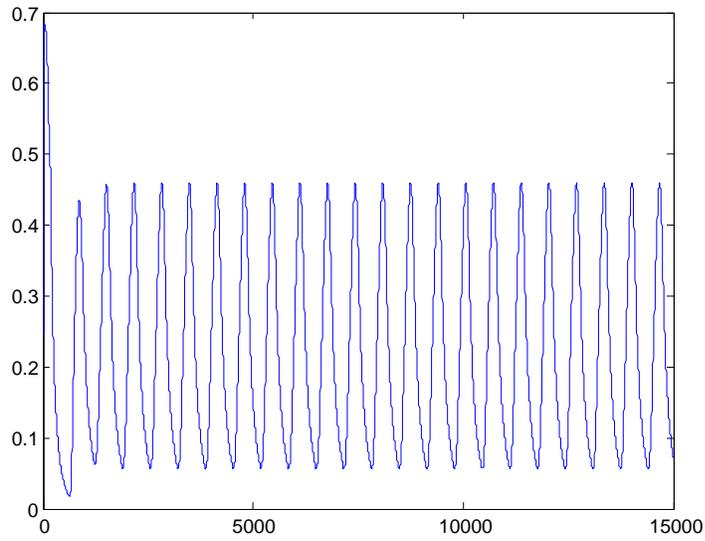


Figure 22: Network Attractiveness as Perceived by the Media, a_m

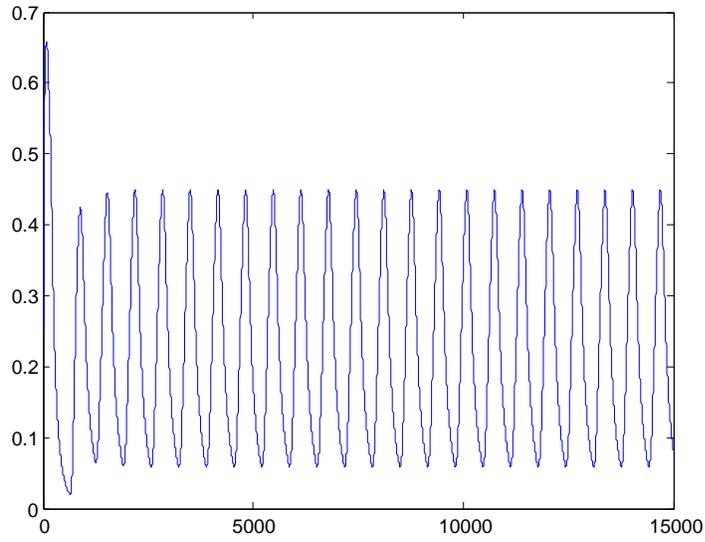


Figure 23: Network Attractiveness as Perceived by Users, r_u

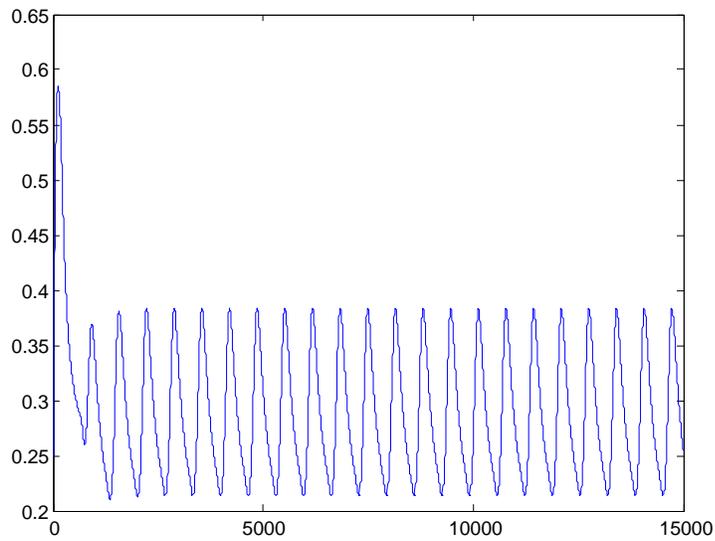


Figure 24: Network Attractiveness as Perceived by Nonusers, r_n

References

- [1] Pavlov, Oleg: *Dynamic Analysis of an Institutional Conflict: Copyright Owners against Online File Sharing*, *Journal of Economic Issues*. 2005. v.39 (3): 633-663
- [2] Pavlov, Oleg and Saeed, Khalid: *A Resource-Based Analysis of Peer-to-Peer Technology*, *System Dynamics Review*. 2004. v.20 (3): 237-262
- [3] Wikimedia Foundation, Inc. "Simulink". April 16, 2006.
<http://en.wikipedia.org/wiki/Simulink>

A Lookup Function Vectors

Latency Acceptance - $f_l(\lambda)$:

Input: [1;2.994;3.813;4.034;4.45;4.694;4.878;4.927;4.988]

Output: [1;.9254;.8202;.7325;.5921;.4035;.2018;.1491;.1]

Content Attractiveness - $f_c(s)$:

Input: [0;0.3823;1.07;2.034;3.364;5]

Output: [0;0.4781;0.7149;0.8991;0.9649;1]

Probability of unsuccessful Download - $p_{ud}(\lambda)$:

Input: [0;.1101;.2141;.2783;.3731;.4648;.5505;.6177;.685;.7431;.8196;.896;1]

Output: [0;.008772;.01754;.03947;.07895;.1404;.2018;.2982;.4605;.7061;.8904;.9649;1]

Delay Factor - $l_{df}(\lambda)$:

Input: [0.06116;1.896;2.661;9.939]

Output: [0.06579;0.4605;0.7895;4.978]

Attrition Response to network attractiveness - $g_d(r_u)$:

Input: [0;.01835;.04587;.104;.1468;.2141;.2997;.3823;.4585;.4985;.6422;.737;.8073;.8685;1]

Output: [10;7.868;5.579;3.408;2.816;2.303;1.908;1.632;1.48;1.43;1.17;1.1;1.05;1.05;1]

Free-riding Multiplier - $f(U)$:

Input: [0;0.1;0.2141;.367;.6116;1.009;2.018;2.966;3.914;5.107;5.872;6.758;7.768;8.502;
9.052;10;10.5;11;11.5;12;12.5;13;13.5;14;14.5;15;15.5;16;16.5;17;17.5;18;18.5;19;19.5;20;20.5;21]

Output: [1;1.3;1.987;2.5;3.803;4.711;6.25;7.197;7.908;8.461;8.895;9.123;9.368;9.605;9.724;
10;10.125;10.25;10.375;10.5;10.625;10.75;10.83;10.89;10.95;11.01;11.06;.9995/.09;
11.10557;11.106;11.107;11.1075;11.108;11.1083;11.1087;11.1091;11.1095;11.1099]

B Simulink Diagram

