

Riikka Susitaival and Samuli Aalto. Modelling the population dynamics and the file availability in a BitTorrent-like P2P system with decreasing peer arrival rate. In Proceedings of the International Workshop on Self-Organizing Systems (IWSOS 2006), pages 34-48, September 2006.

© 2006 Springer Science+Business Media

Reprinted with kind permission of Springer Science+Business Media.

# Modelling the Population Dynamics and the File Availability in a BitTorrent-Like P2P System with Decreasing Peer Arrival Rate

Riikka Susitaival and Samuli Aalto

Helsinki University of Technology  
P.O. Box 3000, FIN-02015 TKK, Finland  
{riikka.susitaival, samuli.aalto}@tkk.fi

**Abstract.** Many measurement studies of P2P file sharing systems suggest that the request rate for a file changes over time and the system is thus non-stationary. For this reason we study the population dynamics and the availability of a file in a BitTorrent-like file sharing system, when the arrival rate for file requests decreases exponentially. We study the system first by a deterministic fluid model and then by a more detailed Markov chain analysis that allows estimating the life time of a single chunk exactly. Simple approximation for the life time is also derived. In addition, we simulate the life time of a file consisting multiple chunks in order to verify the analytical results to be applicable also to a more complex system.

## 1 Introduction

Peer-to-peer (P2P) applications, such as file sharing, have become a significant area of Internet communication in recent years. Older examples of these applications are Gnutella, Napster and Kazaa, whereas BitTorrent is currently the most popular file sharing system. It has been widely reported that P2P related traffic forms a major part of the total traffic in the Internet and the share is even increasing [1], [2].

The idea of BitTorrent is to divide the file into parts, named *chunks*, so that different parts can be downloaded from several peers simultaneously, where the size of the chunk is typically 256 KB. The technical details of BitTorrent are skipped in this paper but can be found in [3]. According to performance studies [4], BitTorrent is an effective P2P protocol and scales well even when the number of participating peers is very high. In this paper we concentrate on BitTorrent-like P2P protocol because of its popularity but the results are applicable to other similar protocols as well.

Measurement studies [5], [6], [7], have shown that the evolution of a single file in the system can be divided into three phases. In the first *flash crowd* phase the demand for the newly released file is high. The flash crowd phase is followed by a *steady state* in which the demand for the file and the service capacity of the system are in balance. Due to the decentralized manner of BitTorrent, there are

not any guarantees that all chunks of the file are present in the system over time. If some chunk of the file is missing, the file is not complete anymore. Depending on application, this might or might not be crucial. In this paper we assume the former. This third phase of the system life time is called as *end phase*.

A few papers have analyzed P2P file sharing systems by stochastic models so far. Yang et al. divide the analysis of BitTorrent-like system into transient and steady state regimes [8]. The service capacity of the transient regime is studied by a branching process and the steady state by a Markov model. Paper [4] studies the performance of the system by a deterministic fluid model, whereas in paper [9] the network level latencies are modeled by the delay of a single class open queueing network and peer level latencies by the delay of M/G/1/K processor sharing queues. However, these early studies do not capture all aforementioned phases of the sharing process. The arrival rate is assumed to be constant, and especially, the file is assumed to be complete forever.

More detailed models for BitTorrent-like P2P systems are provided in [10] and [11]. Tian et al. study the integrity of the file in [10]. First, using Markov chain modelling, the steady state solution for the number of downloaders holding a certain number of chunks is derived and then the file availability is estimated by a simple probabilistic model. Drawback of the paper is that the steady state solution of the system is assumed to exist and the availability model is only indicative. Paper [11] first proposes that arrival rate decreases exponentially and studies then the system by a deterministic fluid model. The life time of the system, i.e., the mean time from the appearance of the file until its disappearance is derived by a simple deduction. However, the model is very coarse, and does not take the download/upload time of the file into account, for example. As a conclusion, the correct analytical models for file availability are missing and should be studied in more detail.

In this paper we analyze the population dynamics of the P2P file sharing system, when the demand for the file decreases exponentially. We use three approaches to study the scenario: deterministic fluid modelling, time-dependent Markov chain modelling and simulations. First we construct a deterministic fluid model for sharing a single piece of the file and study the dynamics of the average number of downloaders and seeds over time by solving the differential equations of the model.

The deterministic fluid model is, however, unable to capture all the details of the chunk sharing process such as possible instability and extinction of the system. For this reason we construct a complete non-homogenous Markov chain model to obtain more information of the life cycle of chunk sharing process. From the Markov chain model we are able to calculate the mean absorption time and thus evaluate the life time of the chunk sharing system more precisely than in papers [10], [11]. We also derive approximations of the life time in some limiting scenarios and evaluate them by simulations using different parameter combinations.

As both the fluid model and the Markov chain model concentrate on the sharing single chunks independently of the other chunks, we study by simulations the

P2P file sharing system with multiple chunks. The obtained analytical results for the life time of a single chunk are compared to simulations of the corresponding system with multiple chunks.

We have studied the population dynamics of the P2P file sharing system already in paper [12]. In the previous paper we assumed that peer requests for the chunk arrive with a constant rate, whereas in this paper we let the peer arrival rate to change over time. Due to this fundamental difference in the assumptions concerning the arrival process, also the deterministic fluid models, Markov models and their solutions characterizing the population dynamics of the system are completely different in the papers.

As a summary, our main contributions in this paper are:

- We construct a deterministic fluid model for chunk sharing and analyze the number of the downloaders and seeds over time.
- We develop a time-dependent Markov chain model for the number of seeds and downloaders and calculate the mean life time of the chunk sharing process. Also approximations are given.
- We verify the applicability of a single-chunk model to estimate the mean life of file sharing by simulating the system with multiple chunks.

The paper is organized as follows: In Section 2 we introduce a model for sharing a chunk with exponentially decreasing arrival rate. Section 3 studies the population dynamics of sharing of a chunk by a deterministic fluid model. In Section 4 we present an accurate result for the mean life time and in Section 5 we approximate it. Section 6 simulates the system with multiple chunks and finally, section 7 makes a short conclusion.

## 2 A Model for Sharing a Single Chunk

In the sections from 2 to 5 the dynamics of the file sharing system is analyzed by concentrating on a single chunk of the file. We study how the number of downloaders and seeds evolves over time from the emergence of the chunk to the disappearance of it. The disappearance of a single chunk means the death of the whole file sharing process since the file is not entire anymore. The work is motivated by the models of [8], [4], [11], but has some differences. In papers [4], [11] the problem of sharing of several chunks concurrently is solved by assuming that peers can forward the chunks with a constant rate. However, we find the assumption unrealistic and the model probably hides some details of the population dynamics. For this reason we consider the sharing of a single chunk at a time. In addition, papers [8] and [4] assume that at least one seed stays in the system keeping the chunks available. However, due to lack of centralized management of BitTorrent, we believe that this assumption is not always true.

In many studies ([8], [4]) the inter-arrival time of new peers is assumed to be exponentially distributed with a constant arrival rate. However, some measurement studies, such as [6] [11], show that the peer arrival rate decreases over time. So in this paper we study the population dynamics of the P2P system with

time-dependent arrival rate  $\lambda(t)$  for a request of the chunk. Paper [11] proposes a traffic model for P2P networks, where the arrival rate decreases exponentially from the release of the file:

$$\lambda(t) = \lambda_0 e^{-\frac{t}{\tau}}, \quad (1)$$

where parameter  $\tau$  describes the attenuation of the demand over time. The average total number of requests arriving in the system, denoted by  $N$ , is then:

$$N = \int_0^{\infty} \lambda_0 e^{-t/\tau} dt = \lambda_0 \tau. \quad (2)$$

According to [11], the exponentially decreasing arrival rate fits well to the corresponding measured peer request rate. For this reason we assume (1) to be a reasonable model to capture the inherent flash crowd effect.

The downloader can download the chunk with rate  $\mu_d$ . On the other hand, the maximum upload rate of a peer for the chunk is assumed to be  $\mu_s$ . After the download, the status of the downloading peer changes from a *downloader* to a *seed*. Note that in this context, a peer is referred to as the seed if it has the chunk in question, but not necessarily all chunks of the file. Seeds stay in the system for downloading other chunks of the file as well as unselfishly serving other peers. All together, the time of staying in the system after downloading the given chunk is assumed to be exponentially distributed with mean  $1/\gamma$ .

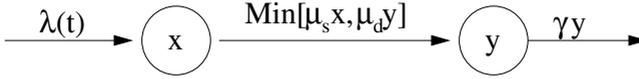
Let  $x(t)$  be the number of downloaders and  $y(t)$  be the number of seeds at time  $t$ . In the next sections we study the evolution of the pair  $(x, y)$  both by a deterministic fluid model and by a Markov model.

### 3 Deterministic Fluid Model

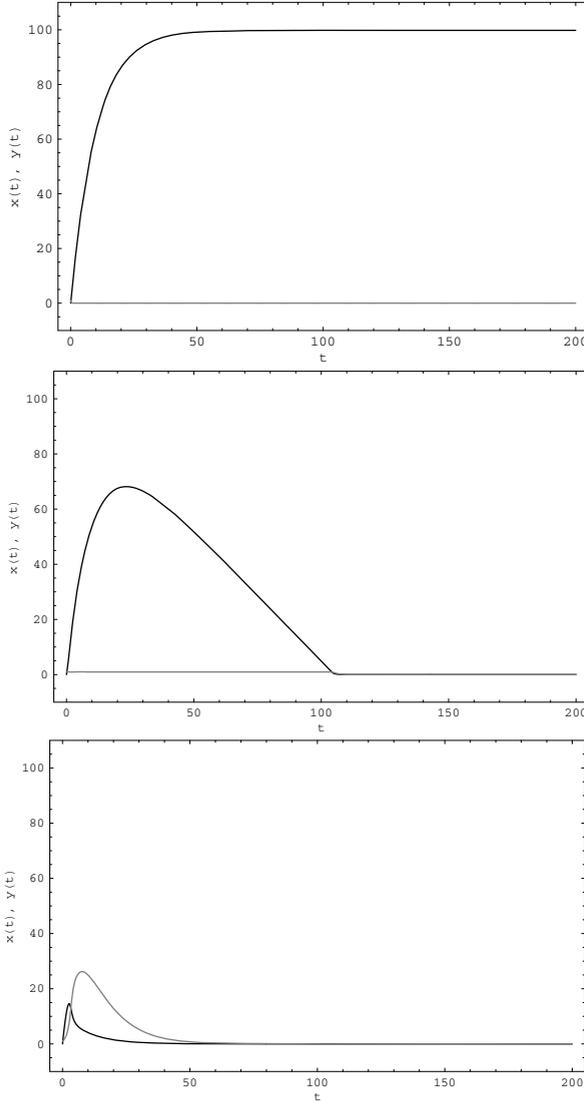
As described in the previous section, we consider a system where a peer starts to spread a single chunk to other peers that are willing to download it. In the system, if the total download capacity of the downloaders,  $\mu_d x(t)$ , is smaller than the total upload capacity of the seeds,  $\mu_s y(t)$ , the downloaders can not use all service capacity provided by the peers. On the other hand, when  $\mu_d x(t) > \mu_s y(t)$  the upload capacity of seeds limits the download process. Thus the total service rate of the system is  $\min\{\mu_d x(t), \mu_s y(t)\}$ . In Figure 1 we have illustrated the model by a flow diagram. The evolution of the number of downloaders and seeds, pair  $(x, y)$ , can be described by a deterministic fluid model. The differential equations of the model are:

$$\begin{aligned} \frac{dx(t)}{dt} &= \lambda_0 e^{-\frac{t}{\tau}} - \min\{\mu_d x(t), \mu_s y(t)\}, \\ \frac{dy(t)}{dt} &= \min\{\mu_d x(t), \mu_s y(t)\} - \gamma y(t), \end{aligned} \quad (3)$$

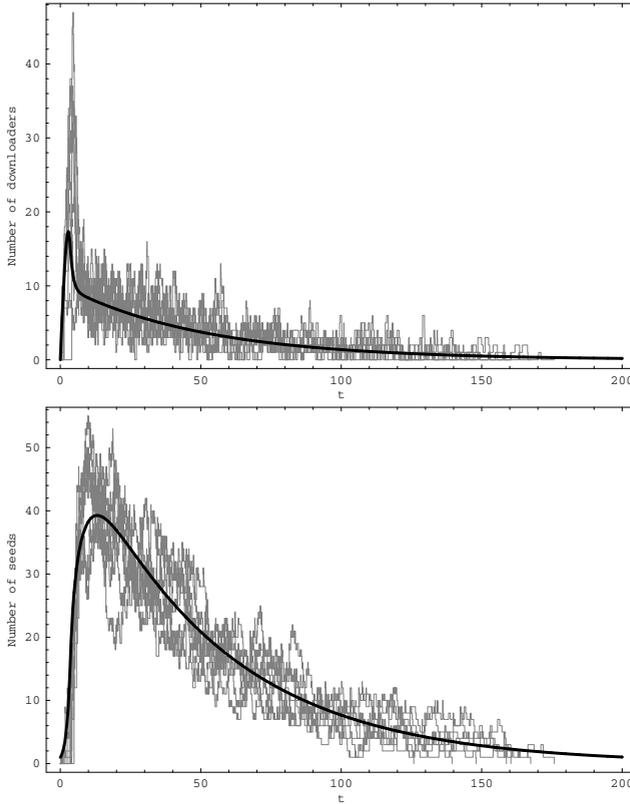
where  $y(0) = 1$  and  $x(0) = 0$  meaning that at the beginning,  $t = 0$ , there is one seed and none downloaders. Let  $\bar{x}$  and  $\bar{y}$  be possible equilibrium values of  $x(t)$  and  $y(t)$ . The differential equations (3) have different solutions depending on the parameters  $\mu_s$ ,  $\mu_d$  and  $\gamma$ . Assuming that the mean download and upload times



**Fig. 1.** A model for sharing a chunk



**Fig. 2.** The number of downloaders and seeds as a function of time, when  $\lambda_0 = 10$ ,  $\tau = 10$ ,  $\mu = 1$  and  $\gamma = 5$  (top),  $\gamma = 1$  (middle) and  $\gamma = 1/5$  (bottom). Black lines: downloaders, gray lines: seeds.



**Fig. 3.** The number of downloaders and seeds as a function of time, when  $\lambda_0 = 10$ ,  $\tau = 50$ ,  $\mu = 1$  and  $\gamma = 1/5$ . Solid line: fluid model, Gray lines: simulation.

are same,  $\mu_s = \mu_d := \mu$ , there are three characteristic solutions for the system, which are depicted by a numerical example of Figure 2:

1.  $\mu < \gamma$ . The existing seeds leave the system faster than new seeds arise. For that reason the number of seeds goes to zero and the number of downloaders increases until all  $N$  requests have arrived (see the top side of Figure 2). The steady-state solution is thus  $\bar{x} \approx N$  and  $\bar{y} = 0$ . Note that these  $N$  downloaders do not receive the chunk at all.
2.  $\mu = \gamma$ . After a very short initial period, the number of seeds stays constant, while the number of downloaders increases beyond that (see the middle of Figure 2). Finally, due to the attenuation of the arrival rate, the number of downloaders returns back to the level of the number of seeds, after which they both go to zero. The steady-state solution is thus  $\bar{x} = 0$  and  $\bar{y} = 0$ .
3.  $\mu > \gamma$ . Both number of the downloaders and seeds increase until all downloaders are served (see the bottom side of Figure 2). Also in this case  $\bar{x} = 0$  and  $\bar{y} = 0$ .

Next we compare the deterministic fluid model to simulations of the system. The number of downloaders and seeds as a function of time is shown in the top and bottom side of Figure 3, respectively. We have fixed  $\lambda_0$  to a moderately small value in order to better demonstrate the dynamics of the system. The black lines correspond to the solutions of equations (3) and gray lines to 10 different simulations of the model described by Figure 1. Simulations have been done by a simple event-based simulator. In the figures we can see that the number of downloaders increases suddenly and the number of seeds increases also a little bit later. As  $t \approx 200$ , both  $x(t)$  and  $y(t)$  are decreased practically to zero. Also all simulation processes have ended due to all seeds had left the system and the chunk is not available anymore.

## 4 Time-Dependent Markov Chain Model for Chunk Sharing

The deterministic fluid model of the previous subsection describes the average behavior of the sharing of a chunk. However, from the simulation results we saw two effects in the population dynamics that were not captured by the fluid model. First, when the chunk became available the earliest seeds could not serve all the downloaders. This capacity shortage was seen as a peak of downloaders during the first 20 time units. Second, if the original seed can leave the system, sooner or later the number of seeds goes to zero. Thus the end of the whole file sharing process is irrevocable. In the simulations all processes were died before  $t = 200$ . The limited life span of the file sharing process has an influence on the performance of the total P2P system and has to be analyzed by some other models than the deterministic fluid model. To this end, in the next subsections we study the evolution of the process  $(x, y)$  in more detail by a non-homogeneous Markov chain model with absorption.

In paper [12] we have derived an analytical model for the life time of the file sharing assuming that the arrival rate of new requests is constant. The result of the paper was that the mean life time of the system increases exponentially as the expected number of the seeds in the system increases. In addition, when the mean downloading time of the chunk was very short, the system can be modelled as an  $M/M/\infty$ -queue with arrival rate  $\lambda$  and departure rate  $\gamma$ . The mean life time equals then to the average length of the busy period  $E[B]$  of  $M/M/\infty$ -queue:

$$E[B] = \frac{1}{\lambda}(e^{\lambda/\gamma} - 1). \quad (4)$$

Contrary to [12], in this paper we assume that peers arrive according to Poisson process with time-dependent rate  $\lambda(t)$ . The mean departure time of a seed,  $1/\gamma$ , the mean download time  $1/\mu_d$  and the mean upload time  $1/\mu_s$  remain constant. Let  $\pi_{(x,y)}(t)$  denote the probability of state  $(x, y)$  at time  $t$ . Next we form the time-dependent Kolmogorov's forward equations for transition rates between the states:

$$\begin{aligned} \frac{d}{dt}\pi_{(x,y)}(t) &= \lambda(t)\pi_{(x-1,y)}(t) + \min\{\mu_d x + 1, \mu_s y - 1\}\pi_{(x+1,y-1)}(t) \\ &\quad + (y + 1)\gamma\pi_{(x,y+1)}(t) - (\lambda(t) + \min\{\mu_d x, \mu_s y\} + \gamma y)\pi_{(x,y)}(t), \end{aligned}$$

for all  $x \geq 1, y \geq 1$ ,

$$\frac{d}{dt}\pi_{(x,0)}(t) = \lambda(t)\pi_{(x-1,0)}(t) + \gamma\pi_{(x,1)}(t) - \lambda(t)\pi_{(x,0)}(t), \text{ for all } x \geq 1, y = 0, \quad (5)$$

$$\begin{aligned} \frac{d}{dt}\pi_{(0,y)}(t) &= \pi_{(\mu_d 1, \mu_s y - 1)}(t) + (y + 1)\gamma\pi_{(0,y+1)}(t) \\ &\quad - (\lambda(t) + \gamma y)\pi_{(0,y)}(t), \text{ for all } x = 0, y \geq 1, \end{aligned}$$

$$\frac{d}{dt}\pi_{(0,0)}(t) = \gamma\pi_{(0,1)}(t) - \lambda(t)\pi_{(0,0)}(t), \text{ when } x = 0, y = 0,$$

with initial condition  $\pi_{(x,y)}(0) = 1$ , when  $x = 0$  and  $y = 1$  and  $\pi_{(x,y)}(0) = 0$  otherwise. The explicit solution of the state probabilities is very complicated but the problem can be solved numerically using a differential equation solver.

In this system all states  $(x, y)$  with  $y = 0$  are absorbing. Let  $z_{(x,y)}$  denote the mean time spent in state  $(x, y)$  from the beginning ( $t = 0$ ) to the absorption of the system:

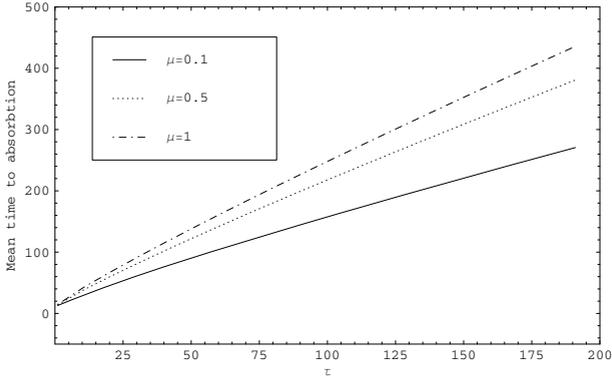
$$z_{(x,y)} = \int_0^{\infty} \pi_{(x,y)}(t) dt. \quad (6)$$

The mean time to absorption, i.e. the life time of the system, is thus the sum of the time spent in the non-absorbing states:

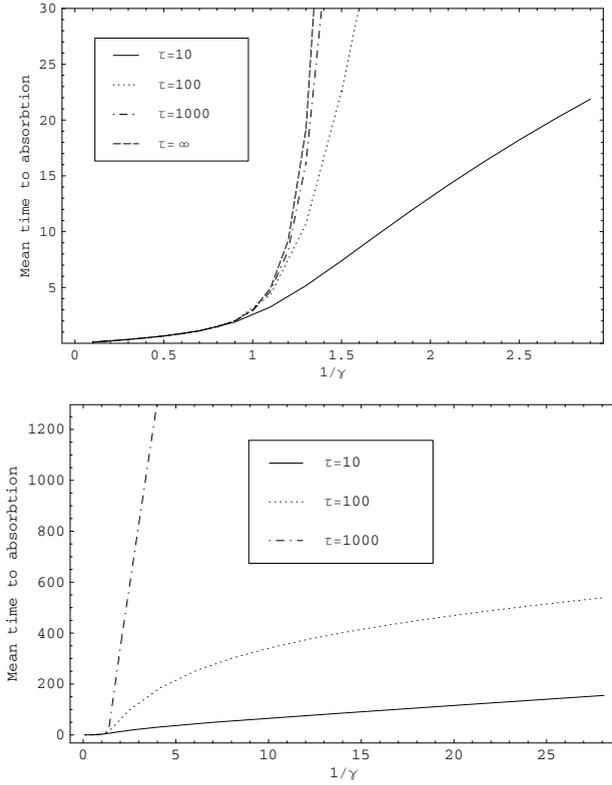
$$T_{life} = \sum_{(x,y):y>0} z_{(x,y)}. \quad (7)$$

By the proposed model above we first study how the life time of the file sharing process depends on the attenuation of the demand. As we increase parameter  $\tau$ , the arrival rate of the requests decreases slower and we have more downloaders in the system. That indicates also more seeds, and a longer life time of the system. This can be seen in Figure 4, which depicts the almost linear growth of the mean life time as a function of  $\tau$  for different download times. For simplicity we assume that the mean upload and download times are the same, denoted by  $1/\mu$ . Also the shorter mean download time increases the life time of the system. The results are obtained from the numerical solution of the mean absorption time (7) when the state space of the Markov process is truncated to  $20 \times 20$  states.

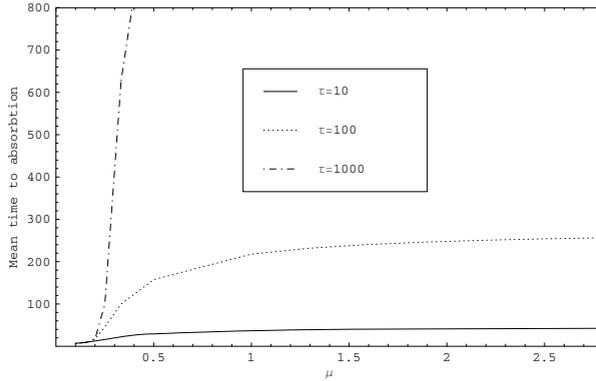
By the model we can also study how the patience of the seeds to stay in the system as servers after own download affects on the availability of the file. In the top of Figure 5 the mean life time of the chunk sharing process is shown as a function of  $1/\gamma$  first for the relatively small values of  $1/\gamma$ . In the figure  $\lambda_0 = 10$  and  $\mu_d = \mu_s = 1$  and  $\tau$  varies from 10 to  $\infty$ . We can see that the mean life time of the system increases exponentially as a function of  $1/\gamma$ . For example, when  $\tau = 1000$ , the mean life time of the system is close to the system with constant arrival rate (case  $\tau = \infty$ ). For smaller  $\tau$ , such as  $\tau = 10$ , the increase is more moderate. The bottom of Figure 5 presents the results for same parameter



**Fig. 4.** The mean life time of the process a function of  $\tau$ . Parameters  $\lambda_0 = 10$ ,  $1/\gamma = 5$ ,  $\mu_s = \mu_d = \mu$  varies from 0.1 to 1.



**Fig. 5.** The mean life time of the chunk sharing process when  $1/\gamma$  varies from 0 to 3 (top) and from 0 to 30 (bottom). Attenuation parameter  $\tau$  varies from  $\tau = 10$  to  $\tau \rightarrow \infty$ . Parameters  $\lambda_0 = 10$  and  $\mu_s = \mu_d = 1$ .



**Fig. 6.** The mean life time of the process as a function of  $\mu$ . Parameters  $\lambda_0 = 10$ ,  $1/\gamma = 5$  and  $\tau$  varies from 10 to 1000.

combination of  $\lambda_0$ ,  $\mu_d$ ,  $\mu_s$  and  $\tau$ , but the mean departure time  $1/\gamma$  is longer. From the results we can see that the ultimate increase of the mean life time as a function of  $1/\gamma$  is only linear when  $\tau < \infty$ .

Last we study the mean life time as a function of the download time of a chunk. Figure 6 depicts the mean life time of the chunk sharing as a function  $\mu$ . When  $\mu \approx 0$ , the mean download time is very long and none of the downloaders has got the file before the original seed has left the system. Thus the mean life time of the system is  $1/\gamma$ . For bigger  $\mu$  the mean life remains almost constant since the system is characterized more by attenuation parameter  $\tau$  and departure rate  $\gamma$ .

## 5 Approximation of the Life Time

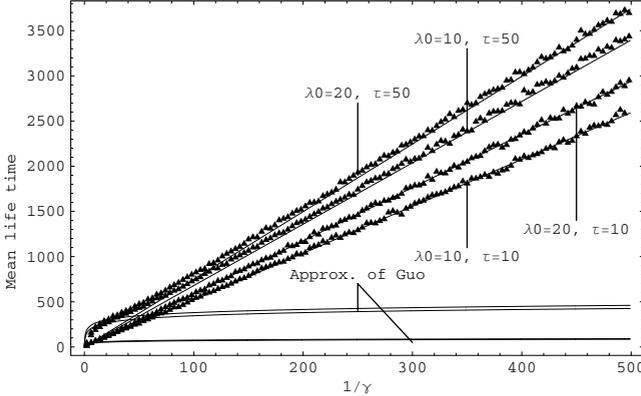
In the previous section we presented an analytical result for the mean life time of sharing a single chunk in the P2P file sharing system. Next we will approximate the life time in a limiting case.

Also paper [11] approximates the mean time to extinction. The simple assumption is that the system dies when  $\lambda(t) < \gamma$ . The mean life time of the system  $T_{life}$  can be then approximated by solving  $t$  from equation  $\lambda(t) = \gamma$ :

$$T_{life} = \tau \log\left(\frac{\lambda_0}{\gamma}\right). \quad (8)$$

However, we will see that this is a very rough approximation for the life span of the bitTorrent-like P2P system and much better bounds for the mean time to absorption can be given.

We consider a case where the mean time that seeds spend in the system is long as compared to the length of the burst of arrivals (meaning the time period



**Fig. 7.** The mean life time of the process as a function of  $1/\gamma$ . First four lines (from bottom): Approximation of paper [11]. Next four lines: our approximation (line) together with corresponding simulations (dots), when a)  $\lambda_0 = 10$  and  $\tau = 10$ , b)  $\lambda_0 = 20$  and  $\tau = 10$ , c)  $\lambda_0 = 10$  and  $\tau = 50$ , d) and  $\lambda_0 = 20$  and  $\tau = 50$ . Parameter  $\mu_d = \mu_c = 1$ .

from the beginning to the moment when  $\lambda(t) \approx 0$ ). If we assume that any seed has not left the system before the end of arrival burst, the life time of the chunk sharing system can be approximated by the two consecutive time periods,  $T_1$  and  $T_2$ , where  $T_1$  is the length of the arrival burst and  $T_2$  is the length of the period from the first to last departure of the seeds:

$$T_{life} \approx T_1 + T_2. \tag{9}$$

First, the length of period  $T_1$  depends on the attenuation of the demand, which is described by parameter  $\tau$ . If we wait fraction  $p$  of total  $N$  arrivals, the length of the burst period is:

$$T_1 = \tau \log\left(\frac{1}{1-p}\right). \tag{10}$$

After the arrival burst all downloaders gradually change their status from the downloader to the seed. We have  $N$  seeds in the system and the mean departure time of all  $N$  seeds is

$$T_2 = \frac{1}{N\gamma} + \frac{1}{(N-1)\gamma} + \frac{1}{(N-2)\gamma} \dots + \frac{1}{\gamma} = \sum_{i=0}^{[N]} \frac{1}{i\gamma}. \tag{11}$$

Next we assume that the mean departure time is long as compared to the attenuation of the demand,  $1/\gamma \gg \tau$ . In this case also  $T_2 \gg T_1$  and we can approximate  $T_{life} \approx T_2$ . In Figure 7 we compare the proposed approximation to the simulation of the corresponding system with different combinations of

$\lambda_0$  and  $\tau$ . If  $1/\gamma$  is long, there are lot of seeds in the system and we use simulation instead of analytical model due to limited state space of the Markov model (simulation results are averaged over 100 simulation replications). Also the approximation of paper [11] is presented in Figure 7. For small values of  $1/\gamma$  neither of the approximations are good. However, as  $1/\gamma > \tau$  we can see that our approximation is very accurate for many parameter combinations as compared to the life time proposed by paper [11].

## 6 Sharing of Multiple Chunks

In the previous sections we considered sharing of a single chunk as an independent process of other chunks and estimated the life time of the file sharing system by the disappearance of a single chunk. However, in reality the dynamics of a single chunk depends on the other chunks as well. So in this section we consider population dynamics of a more complex system with multiple chunks.

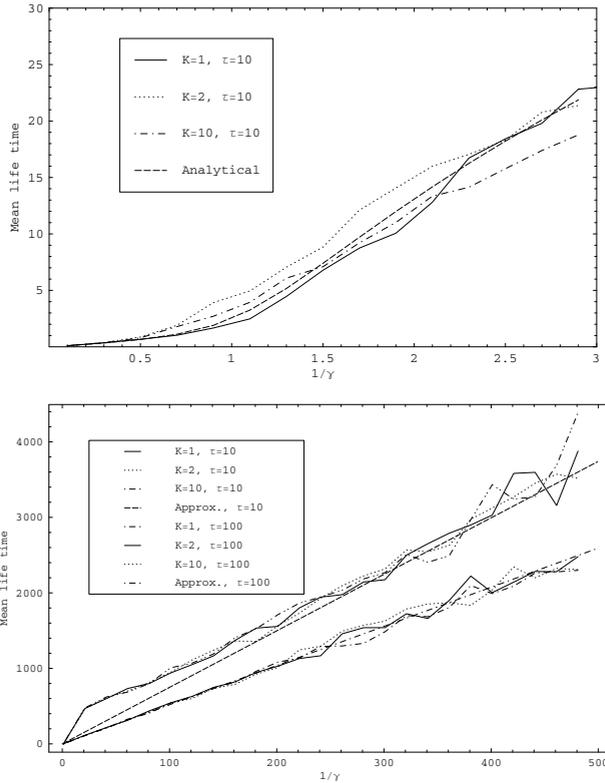
In the system we have three types of peers, *downloaders*, *leechers*, and *seeds*. Downloaders do not have any chunk yet and try to only download a first one. Leechers have already downloaded a part of the chunks and can upload those chunks to other peers in the system. While uploading the chunks, they try to find the rest of the chunks. Finally, peers that have all chunks are called seeds. These seeds stay for a while in the system uploading the chunks to the downloaders.

Let  $L$  be the size of the file under consideration in bytes. The file is divided into blocks, named chunks, which are assumed to be downloaded one after another separately. If we fix the size of chunk to  $l$  bytes, the total number of chunks is naturally  $L/l$ , denoted by  $K$ . New peers requesting the given file are assumed to arrive at the system with time-dependent rate  $\lambda(t)$ . Also in this section we assume that request rate decreases exponentially according to formula (1).

When new downloader  $i$  has arrived, it seeks random peer  $j$  among all available peers including leechers and seeds in the system and compare the own set of chunks and peer  $j$ 's set of chunks. If downloader  $i$  finds a chunk that is not in its own collection, it starts to download it. We fix the mean download time of a chunk to be proportional to the chunk size, that is,  $\frac{1}{K\mu_d}$ . Also the maximum upload rate of a peer is proportional to the chunk size,  $K\mu_c$ . If there are many downloaders per one leecher or seed, the capacity of the peer is divided among downloaders. We assume that time required to find the peer for exchanging chunks is negligible as compared to the download time.

After the peer has downloaded the first chunk, it seeks a next random peer and a next random chunk. Also if selected peer  $j$  has not any new chunk that the downloader  $i$  does not already have, downloader seeks a new peer. When a peer has some, it can as a leecher upload the chunk to other peers, if required. When the leecher has collected all  $K$  chunks, the status of the peer changes from a leecher to a seed. We assume that the seeds stay in the system for a random period. Let  $\gamma$  denote the departure rate of a seed.

However, we do not assume that there is always a seed keeping the file entire in the system. If some of the chunks is missing, i.e. there are no seeds in the system



**Fig. 8.** The mean life time of the process as a function of  $1/\gamma$ , when the file is divided into multiple chunks. Number of chunks varies from  $K = 1$  to  $K = 10$  and  $\tau$  varies from  $\tau = 10$  to  $\tau = 100$ . Parameters  $\lambda_0 = 10$ ,  $\mu_d = \mu_c = 1$ .

and the chunk is not included in any chunk collections held by the leechers, the file is not complete anymore. In this case the sharing process of the given file dies.

First in the top of Figure 8 we have simulated the file sharing process for small values of  $1/\gamma$ , when the number of chunks varies from  $K = 1$  to  $K = 10$  and  $\tau = 10$ . The number of the simulation replications for a given value of  $1/\gamma$  is 500. From the figure we can see that our analytical model for sharing of a single chunk corresponds well to the system with multiple chunks.

In the bottom of Figure 8 we have simulated the same file sharing process for big values of  $1/\gamma$ . Lower three lines correspond to  $\tau = 10$ , and upper three lines to  $\tau = 100$ . Also the approximations for one chunk of the previous section are presented. The result is that also in these cases the mean life time of the system increases linearly as a function of  $1/\gamma$  and fits well to approximation of the chunk model presented in Section 5. For these parameter combinations, the selected coarseness of the division of the file into chunks does not influence substantially the mean life time of the file sharing process.

As a result of this section we can say that the mean life time calculated from the one chunk model correspond well to the life time of the more realistic system with multiple chunks. This information helps in performance evaluation and optimization of the file sharing systems with varying system parameters.

## 7 Conclusion

In this paper we have studied the population dynamics and the file availability of the BitTorrent-like P2P system. Our approach was to calculate the mean life time of a single chunk independently of the other chunks. We considered both a deterministic fluid model and a non-homogenous Markov chain model. By the latter model we were able to study how the different parameters, such as the attenuation of the demand, the mean download time and the mean departure time affected the mean life time of the sharing process. We provided also a simple approximation for the life time, which gave better results than an earlier proposal. The applicability of the one chunk model was verified by the simulations of sharing of the file in multiple chunks. Our approach in this paper was rather in modelling the existing systems than in optimization of the P2P file sharing algorithms.

In future we will study more the analytical models and especially approximations of the life time also in other scenarios than we have done so far. For example, the behavior of the system, when seeds can return back to the system, is interesting. In addition, we will deepen the study of the scenario with multiple chunks by both analytical model and more complex simulations.

## References

1. T. Karagiannis, A. Broido, N. Brownlee, kc claffy, M. Faloutsos, Is P2P dying or just hiding?, in Globecom, 2004.
2. <http://www.cachelogic.com/p2p/p2ptraffic.php>
3. B. Cohen, Incentives Build Robustness in BitTorrent, 2003, in Proc. of First Workshop on Economics of Peer-to-Peer Systems, June 2003, <http://www.bittorrent.com/bittorrentecon.pdf>.
4. D. Qiu, R. Srikant, Modeling and Performance Analysis of BitTorrent-Like Peer-to-Peer Networks, in Proc. of SIGCOMM 2004.
5. M. Izal, G. Uvroy-Keller, E.W. Biersack, P.A. Felber, A.Al Hamra, and L. Garcés-Erice, Dissecting BitTorrent: Five Months in a Torrent's Lifetime, in Proc. of PAM, 2004.
6. J.A. Pouwelsem P. Garbacki, D.H.J. Epema, H.J. Sips, The BitTorrent P2P File-sharing system: Measurements and analysis, in Proc. of IPTPS, 2005.
7. L. Massoulié and M. Vojnović, Coupon replication Systems, in Proc. of SIGMETRICS, 2005.
8. X. Yang, G. de Veciana, Service Capacity of Peer to Peer Networks, in Proc. of INFOCOM 2004.
9. K.K. Ramachandran, B. Sikdar, An Analytic Framework for Modeling Peer to Peer Networks, in Proc. of INFOCOM 2005.

10. Y. Tian, D. Wu, K. Wing Ng, Modeling, Analysis and Improvement for BitTorrent-Like File Sharing Networks, in Proc. of INFOCOM 2006.
11. L. Guo, S. Chen, Z. Xiao, E. Tan, X. Ding, X. Zhang, Measurement, Analysis, and Modeling of bitTorrent-like Systems, in Proc of Internet Measurement Conference of USENIX Association, 2005.
12. R. Susitaival, S. Aalto, J. Virtamo, Analyzing the dynamics and resource usage of P2P file sharing by a spatio-temporal model, in Proc. of P2P-HPCS'06 in conjunction with ICCS'06, pp. 420–427, 2006.