Chapter 14
Modeling of P2P File Sharing with a Level-Dependent QBD Process

Sophie Hautphenne, Kenji Leibnitz, and Marie-Ange Remiche

Abstract In this chapter we propose to analyze a peer-to-peer (P2P) file sharing system by means of a so-called level-dependent Quasi Birth-and-Death (QBD) process. We consider the dissemination of a single file consisting of different segments and include a model for the upload queue management mechanism with peers competing for bandwidth. By applying an efficient matrix-analytic algorithm we evaluate the performance of P2P file diffusion in terms of the corresponding extinction probability, that is, the probability that the sharing process ends.

14.1 Introduction

With the introduction of peer-to-peer (P2P) technology in networks for file sharing and content distribution, the volume of transported traffic has recently enormously increased. The nodes participating in the P2P network are called peers and form logical overlay structures on the application layer above the IP topology; see Fig. 14.1. One of the main advantages of using P2P networks for content distribution is their high scalability to a growing number of file requests, especially in the presence of flash crowd arrivals [1]. Unlike conventional client/server architectures, all peers act simultaneously as clients and servers, thus shifting the load from a single server to

S. Hautphenne
Département d'Informatique, Université Libre de Bruxelles, B-1050 Bruxelles, Belgium
e-mail: shautphe@ulb.ac.be

M.-A. Remiche
Faculté des Sciences Appliquées, Université Libre de Bruxelles, B-1050 Bruxelles, Belgium
e-mail: mremiche@ulb.ac.be

K. Leibnitz
Graduate School of Information Science and Technology, Osaka University, Osaka 565-0871, Japan
e-mail: leibnitz@ist.osaka-u.ac.jp

W. Yue et al. (eds.), *Advances in Queueing Theory and Network Applications,*

© Springer Science+Business Media LLC 2009

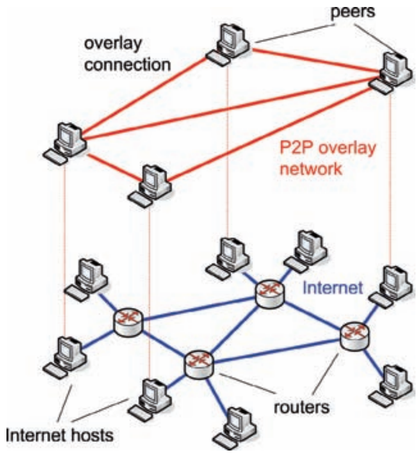

Fig. 14.1 A P2P network consists of peers forming a logical overlay network above the IP topology.

several peers sharing a specific file. Additionally, because the source of a file is no longer stored at a single location, the P2P network is more robust to failures.

However, there are also certain dangers in entirely relying on P2P networks for file distribution. Firstly, as the data are no longer kept at a single trusted source, each peer that hosts the file may modify the data willingly or unwillingly, thus causing the distribution of corrupt information. This is referred to as *poisoning* or *pollution* [2]. Secondly, the existence of a sharing peer in the network cannot be guaranteed due to *churn* (i.e., the process of peers entering and leaving the network). The sharing of files is controlled by the peers' behavior (willingness to share after downloading, patience, etc.) and they may arbitrarily join or leave the network at any instant [3]. If the peer, which has the last part of the file, leaves the network, this information is lost and other peers can no longer retrieve the data. For this reason, specific P2P architectures (e.g., Chord [4]) employ mechanisms to maintain a certain number of replicas of a file in the network.

In this chapter we study the probability that the diffusion of a file will eventually come to a halt in an unstructured P2P file sharing network, which we define as the *extinction* of the file. We extend our previous model in [5], where we used a Markovian Binary Tree (MBT) to model the file sharing network and we formulated an algorithm to compute the extinction probability. However, the previous model only considered the sharing of entire files. In this chapter, we extend the model to include the sharing of individual parts of the file to reflect a more accurate behavior. This is achieved by using a level-dependent Quasi Birth-and-Death (QBD) process. By adapting the logarithmic-reduction algorithm (see Latouche and Ramaswami [6]), we actually compute the probability that file diffusion ends due to the lack of peers sharing a part of the file.

This chapter is organized as follows. First, we briefly summarize some related work on modeling of P2P file sharing mechanisms for content distribution in Sect. 14.2. This is followed by the formulation of our basic assumptions on the file sharing network in Sect. 14.3. Although we consider a P2P network that roughly resembles the eDonkey protocol, the model is general enough to be easily applied to other file sharing protocols as well. In Sect. 14.4 we formulate two analytical models corresponding to two different systems in which either the sharing process stops when the entire file is lost or when any of the segments is missing. Accordingly, we construct the corresponding level-dependent QBD process and we develop algorithms necessary to obtain the extinction probability in both settings. We provide some numerical results showing the impact of the system parameters on the performance of the system in Sect. 14.5. Finally, conclusions are drawn in Sect. 14.6.

14.2 Related Work

A growing number of studies can be found dealing with the modeling and performance evaluation of P2P file sharing networks. In this section we only highlight a few of them that we consider relevant to this chapter. Most studies on the evaluation of P2P systems as content distribution networks rely on measurements or simulations of existing P2P networks. For example, Saroiu et al. [7] conducted measurement studies of content delivery systems that were accessed by the University of Washington. The authors distinguish between traffic from P2P, WWW, and the Akamai content distribution network, and they found that the majority of volume is transported over P2P. In [8], a measurement-based traffic profile of the eDonkey network is provided and reveals that there is a strong distinction between download flows and nondownload streams. Similar studies exist for the Gnutella network [9] and BitTorrent [10], as well. Hoßfeld et al. [11] provide a simulation study of the eDonkey network and examine the file diffusion properties under constant and flash crowd arrivals.

An analytical model for performance evaluation of a generalized P2P system is given by Ge et al. [12]. On the other hand, other published work mostly considers specific existing applications. For example, Qiu and Srikant [13] used a fluid model for BitTorrent and investigate the performance in steady state. They studied the effectiveness of the incentive mechanism in BitTorrent and proved the existence of a Nash equilibrium. Rubenstein and Sahu [1] mathematically showed that unstructured P2P networks have good scalability and are well suited to cope with flash crowd arrivals. A fluid-diffusive P2P model from statistical physics is presented by Carofiglio et al. [14]. Both the user and the content dynamics are included, but this is only done on the file level and without pollution. All these studies show that by providing incentives to the peers for sharing a file, the diffusion properties are improved. Yang and de Veciana [15] investigated the service capacity of P2P networks by considering two models, one for the transient state with flash crowds and one in steady state.

Christin, Weigend and Chuang [2] measured content availability of popular P2P file sharing networks and used these measurement data for simulating different pollution and poisoning strategies. They show that only a small number of fake peers can seriously affect the user's perception of content availability. In [16], a diffusion model for modeling eDonkey-like P2P networks is presented based on a model from mathematical biology. This model includes pollution and a patience threshold at which a peer aborts its download attempt and retries again later. It is shown that an evaluation of the diffusion process is not accurate enough when steady state is assumed or the model only considers the transmission of the complete file, especially in the presence of flash crowd arrivals. That model is extended in [17] to analytically compare the performance of P2P file sharing networks to that of client/server systems.

14.3 Peer-to-Peer File Sharing Model

Let us now define the assumptions we make on the P2P file sharing model in this chapter. We assume an unstructured P2P network operating similar to the eDonkey network. However, our model is not restricted to eDonkey, but can in fact be applied to other file sharing networks as well. The sharing of a file with size F is performed in units of chunks, which are further split into smaller units called blocks; see Fig. 14.2. In eDonkey, a chunk has the size of 9.28 MB and a block is 180 kB. After each chunk has been downloaded, it is checked for errors and if the hash value is incorrect, all blocks of the chunk are discarded and downloaded again. After all chunks of a file have been successfully downloaded, the peer may decide to keep the file as a seeder in the network for other peers to download or to remove the file from sharing (leecher or free rider). In this work, we assume that the file consists only of a single chunk, corresponding, for example, to a single mp3 audio file, as this is enough to capture the basic characteristics of the diffusion behavior.

14.3.1 Upload Queue Management and File Segmentation

In order to manage the bandwidth for other peers requesting the file, an upload queue mechanism is maintained. A peer requests individual blocks from other peers

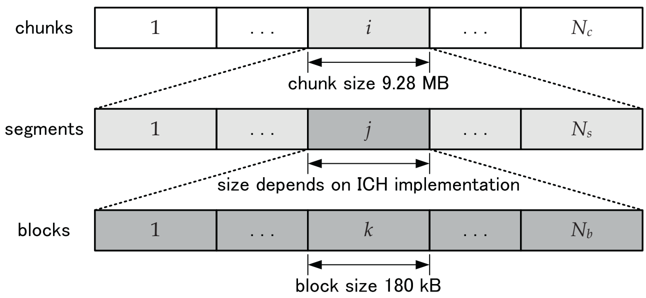

Fig. 14.2 File structure consisting of chunks, segments, and blocks.

sharing the chunk that contains the desired block. All requests are appended to the waiting list of the sharing peer and a weighting mechanism handles the scheduling of the upload queue requests for transmission. The detailed procedure of the queue management takes several features into account that depend on the individual settings of the sharing peer such as upload bandwidth and number of simultaneous uploads.

In our model, an approximative assumption simplifies the upload queue management behavior [11]. If a peer downloads a block from another peer, additional blocks might be of interest, if the providing peer is not already sharing the complete file. The weighting mechanism takes this into account by giving higher priority to peers from which blocks had been previously downloaded. We include this interaction by considering that not individual blocks, but rather a series of blocks is downloaded at a time after moving from the waiting list to the uploading list. The waiting list is modeled as a FIFO (first-in-first-out) queue and the number of consecutively downloaded blocks can be obtained from measurements [8] through the average data volume downloaded per sharing peer.

In the original version of eDonkey, error detection is done after all blocks of a chunk have been received and the complete chunk is discarded in the case of an error. However, this is not very effective and in more recent versions of eDonkey clients (e.g., eMule), the Intelligent Corruption Handling (ICH) mechanism is implemented which performs an error detection on smaller data units than chunks and that we define in the following as segments. Instead of discarding the complete chunk when at least one corrupted block is received, only all blocks of the damaged segment need to be requested again. The actual size of a segment depends on the specific settings of the ICH mechanism.

With the assumptions on the upload queue mechanism and corruption handling, it is sufficient to consider that a chunk only needs to be modeled consisting of few segments instead of several individual blocks. In this study we assume that a chunk consists of two segments (i.e., $N_s = 2$) and the size of a segment is $Z = 4.64$ MB. The size of the whole file F is less than or equal to 9.28 MB.

14.3.2 Download Bandwidth

Let us define the upload and download rates as r_u and r_d, respectively. For the sake of simplicity, we use the same assumption as in [16] of homogeneous users with ADSL connections, resulting in rates of $r_u = 128$ kbps and $r_d = 768$ kbps. Furthermore, let us denote the number of peers sharing a certain segment as S and the peers downloading it as D. Because eDonkey employs a fair share mechanism for the upload rates, there are on average S/D sharing peers serving a single downloading peer and we multiply this value with r_u. This gives us the bandwidth on the uplink.

However, because the download bandwidth could be the limiting factor, the effective downloading rate of a segment consists of the minimum of both terms, that is, $\min(S/Dr_u, r_d)$. When downloading a segment of size Z, the term $\min(S/Dr_u, r_d)/Z$ represents the proportion of the segment that is downloaded in one unit of time, thus, the rate at which we may observe the arrival of new peers that have completely downloaded the file. We call this rate the effective transition rate. It is worth noting that in general the effective downloading rate depends on the interaction of the peers within the system (namely the number of downloaders and the number of peers sharing the segment) and on the size of segment that is effectively downloaded.

14.4 Analytical P2P File Sharing Model

Let us consider a chunk to be made up of two segments: segment 1 and segment 2 of respective sizes Z and $F-Z$, where F, as defined earlier, is the size of the complete file. We end up with three categories of peers; namely, peers with segment 1 or 2 and peers that have both segments. We say that a peer is in phase i ($i = 1,2$) when it possesses only segment i and in phase 3 in the case where it has both segments. New peers are assumed to appear at random times in the system determined by an exponential random variable whose rate depends both on the effective transition rate we introduced above and on the current state of the system, that is, the number of peers S_i in each phase $i = 1,2$, or 3. For the sake of simplicity, we can assume that the rate at which a peer stops sharing a segment is independent of the segment number, and is equal to d. The ensuing model is now described.

Let us now define the stochastic process $\{(X(t), \varphi(t))\}$, where $X(t)$ counts the total number of peers present in the system at time t, and $\varphi(t) = (\varphi_1(t), \varphi_2(t), \varphi_3(t))$ denotes the number of peers in each phase present in the system at time t, with $\varphi(t)\mathbf{1} = X(t)$. Here, $\mathbf{1}$ denotes a vector with ones.

We consider two views to measure the extinction probability of the file sharing process, an optimistic and a pessimistic view. In the optimistic view, we assume that the sharing process ends when no more segments are available in the system. In the pessimistic case, the file sharing process ends as soon as one of the two segments is missing. We call the latter event a catastrophe. Let us explain each resulting model in turn.

14.4.1 Level-Dependent QBD

In this first setting, recall that the sharing process ends when there are no more segments available in the system. The stochastic process $\{(X(t),\ \varphi(t))\}$ is an absorbing level-dependent quasi birth-and-death process, of which the generator Q can be written as

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ A_2^{(1)} & A_1^{(1)} & A_0^{(1)} & 0 & 0 & 0 & \cdots \\ 0 & A_2^{(2)} & A_1^{(2)} & A_0^{(2)} & 0 & 0 & \cdots \\ 0 & 0 & A_2^{(3)} & A_1^{(3)} & A_0^{(3)} & 0 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{14.1}$$

This process has been extensively studied in the past (see Latouche and Ramaswami [6] and references therein). In this setting, the time to extinction of the system is clearly equal to the time until absorption. In the remainder of this section, we first elaborate on the content of the $A_i^{(j)}$ matrices (with $i = 0,1,2$ and $j \geq 1$) and then give the algorithmic procedure in order to compute the absorption probability in this level-dependent QBD with generator Q.

14.4.1.1 Level-Dependent QBD Generator Description

When the system is in state (S_1,S_2,S_3), it means that we have S_1 peers in phase 1 (with only segment 1), S_2 peers in phase 2 (with only segment 2), and S_3 peers in phase 3 (with the complete file). We define the state subspace $L(k)$, $k \in \mathbb{N}$, as

$$L(k) = \{(S_1,S_2,S_3) : S_1 \geq 0, S_2 \geq 0, S_3 \geq 0;\, S_1+S_2+S_3 = k\},$$

which gives all states of the system at level k, that is, when k peers are present in the system. Its cardinality is clearly

$$|L(k)| = \frac{1}{2}(k+2)(k+1)$$

and we take the lexicographic order to enumerate the states of each level.

Before proceeding with the description of the transition matrix, we define two functions of crucial interest in the following; these are

$$\mu_i(S,D) = \frac{1}{Z_i}\min\left\{\frac{S}{D}r_u, r_d\right\}, \qquad i = 1,2, \tag{14.2}$$

where $Z_1 = Z$ and $Z_2 = F - Z$ are the sizes of each segment and S and D are the number of all peers currently sharing and downloading the segment, respectively.

When the system contains a single peer (i.e., when its state is in $L(1)$), this peer may stop sharing the one segment it possesses with rate d (the system then moves to $L(0)$) or another peer may start downloading the segment (the system is thus in $L(2)$). The first event occurs at a rate recorded by $A_2^{(1)}$; that is,

$$A_2^{(1)} = \begin{bmatrix} d \\ d \\ d \end{bmatrix}.$$

The latter case occurs at a rate given by the matrix $A_0^{(1)}$ as

$$A_0^{(1)} = \begin{bmatrix} \mu_1(1,1) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu_2(1,1) & 0 & 0 \\ 0 & 0 & \mu_1(1,1) & 0 & \mu_2(1,1) & 0 \end{bmatrix}.$$

Indeed, if the system is in state $(0,1,0)$, for example, only a new peer with segment 2 may appear; that is, the system moves towards state $(0,2,0)$. This happens at a rate $\mu_2(1,1)$; see (14.2).

Usually, a peer may also perform a change of phase, that is, from 1 to 3 or from 2 to 3. Such a transition keeps the level at 1 because no new peer arrives in the system. However, if a peer in phase 1 (or phase 2) is alone in the system, it will not be able to download the missing segment and to change into phase 3. Thus, the transition rate from phase 1 (or from phase 2) to phase 3 when the system is in level 0, is $\mu_i(0,1) = 0$ for $i = 1,2$ in that particular case. The diagonal elements of $A_1^{(1)}$ (and of all $A_1^{(k)}$, $k \geq 2$) are such that $Q\mathbf{1} = \mathbf{0}$. It finally gives

$$A_1^{(1)} = \begin{bmatrix} -d-\mu_1(1,1) & 0 & 0 \\ 0 & -d-\mu_2(1,1) & 0 \\ 0 & 0 & -d-\mu_1(1,1)-\mu_2(1,1) \end{bmatrix}.$$

The possible transitions from a state $(S_1,S_2,S_3) \in L(k)$ with $k \geq 2$ are described below.

$A_2^{(k)}$: This matrix records the rate at which the system may lose a peer. A peer in phase i disappears with rate d. This latter is multiplied by the number of peers in phase i, that is, S_i with $i = 1,2,3$.

$A_0^{(k)}$: This matrix explains at which average rate a new peer may arrive in the system. There exist two possible transitions, listed in the table below. They both may be interpreted with a similar argumentation, so we limit our explanation to only the first case of possible transitions. The effective downloading rate of a new peer with segment 1 is determined as usual as the minimum between its own physical downloading rate r_d and a rate which depends on the number of peers that are sharing the available total upload bandwidth. Segment 1 is available to peers in phases 1 and 3. However, although there are only S_2 peers interested in downloading segment 1 from peers in phase 1, there are S_1+S_2 peers interested in downloading segment 1 or segment 2 from peers in phase 3. It is important to take into account the S_1 supplementary peers because they also share the available upload bandwidth at peers in phase 3. This leads to an effective transition rate of

$$\mu_3(S_1,S_2,S_3) = \frac{1}{Z}\min\left\{\left(\frac{S_1}{S_2+1}+\frac{S_3}{S_1+S_2+1}\right)r_u, r_d\right\}$$

Table 14.1 Transitions and rates for matrix $A_0^{(k)}$.

Transitions	Rates
$(S_1,S_2,S_3) \to (S_1+1,S_2,S_3)$	$\mu_3(S_1,S_2,S_3)$
$(S_1,S_2,S_3) \to (S_1,S_2+1,S_3)$	$\mu_4(S_1,S_2,S_3)$

Table 14.2 Transitions and rates for matrix $A_1^{(k)}$.

Transitions	Rates
$(S_1,S_2,S_3) \to (S_1-1,S_2,S_3+1)$	$\mu_2(S_2+S_3,S_1)$
$(S_1,S_2,S_3) \to (S_1,S_2-1,S_3+1)$	$\mu_1(S_1+S_3,S_2)$
Diagonal element	Parameter of the exponential
$(S_1,S_2,S_3) \to (S_1,S_2,S_3)$	$-kd-\mu_3(S_1,S_2,S_3)-\mu_4(S_1,S_2,S_3)$ $-\mu_2(S_2+S_3,S_1)-\mu_1(S_1+S_3,S_2)$

and accordingly to

$$\mu_4(S_1,S_2,S_3) = \frac{1}{F-Z}\min\left\{\left(\frac{S_2}{S_1+1}+\frac{S_3}{S_1+S_2+1}\right)r_u, r_d\right\}$$

for the case of a new peer appearing in phase 2. Table 14.1 summarizes the transitions and their corresponding rates.

$A_1^{(k)}$: A peer in phase 1 turns into a peer in phase 3 with the rate $\mu_2(S_2+S_3,S_1)$, because S_1 peers are competing for the $(S_2+S_3)\,r_u$ available bandwidth. The same argument holds for a peer in phase 2 changing into a peer in phase 3. Let us recall that the diagonal elements are such that $Q\mathbf{1}=\mathbf{0}$. The corresponding transitions and rates are shown in Table 14.2.

14.4.1.2 Probability of Extinction

Our interest lies in computing the probability that the sharing process in the particular system setting described in the previous section will terminate at some point. Let $\gamma(0)$ be the first time the system is in level 0; that is no segment is available. Let $\mathbf{e}_i$ be a unit vector with a 1 at the ith entry and 0 elsewhere. In this chapter, an empty product is, by convention, equal to the identity matrix (for $l=0$ in (14.3), for instance). We define $(\mathbf{G}_1)_i$ as the probability that the system starting in level 1 with $\varphi(0)=\mathbf{e}_i$ will eventually reach level 0; that, is,

$$(\mathbf{G}_1)_i = P[\gamma(0)<\infty\,|\,\varphi(0)=\mathbf{e}_i] \qquad i=1,2,3.$$

It was proven in [19] that this vector is explicitly given by

$$\mathbf{G}_1 = \sum_{l=0}^{\infty}\left[\prod_{i=0}^{l-1}U_{2^i}^{i}\right]\mathbf{D}_{2^l}^{l}, \tag{14.3}$$

where

$$U^i_{2^i} = P\left[\gamma(2^{i+1}) < \gamma(0) \wedge \varphi(\gamma(2^{i+1}))\,|\,X(0) = 2^i\right],$$
$$\mathbf{D}^l_{2^l} = P\left[\gamma(0) < \gamma(2^{l+1}) \wedge \varphi(\gamma(0))\,|\,X(0) = 2^l\right]$$

and where $\gamma(k)$ is defined as the first passage time to level k; that is,

$$\gamma(k) = \inf\{t \geq 0 : X(t) = k\}$$

with $k \geq 0$. Accordingly, we have

$$\left[\prod_{i=0}^{l-1} U^i_{2^i}\right] \mathbf{D}^l_{2^l} = P[\gamma(2^l) < \gamma(0) < \gamma(2^{l+1}) \wedge \varphi(\gamma(0))\,|\,X(0) = 1], \tag{14.4}$$

that is, the probability that the process starting from level 1, first visits level 2^l, then visits level 0 before visiting level 2^{l+1}. Summing (14.4) over $l = 0$ to infinity clearly gives $\mathbf{G}_1$.

The matrices U^l_k and D^l_k, respectively, of dimensions $|L(k)| \times |L(k+2^l)|$ and $|L(k)| \times |L(k-2^l)|$, are given by the following recursive equations:

$$U^0_k = \left(-A^{(k)}_1\right)^{-1} A^{(k)}_0, \tag{14.5}$$

$$D^0_k = \left(-A^{(k)}_1\right)^{-1} A^{(k)}_2, \tag{14.6}$$

$$U^l_k = \left[I - U^{l-1}_k D^{l-1}_{k+2^{l-1}} - D^{l-1}_k U^{l-1}_{k-2^{l-1}}\right]^{-1} U^{l-1}_k U^{l-1}_{k+2^{l-1}}, \qquad l \geq 1, \tag{14.7}$$

$$D^l_k = \left[I - U^{l-1}_k D^{l-1}_{k+2^{l-1}} - D^{l-1}_k U^{l-1}_{k-2^{l-1}}\right]^{-1} D^{l-1}_k D^{l-1}_{k-2^{l-1}}, \qquad l \geq 1. \tag{14.8}$$

Note that for $k = 2^l$ the matrix D^l_k will become a vector. A clear proof is given in [19]. The sum in (14.3) needs to be truncated in order to numerically evaluate $\mathbf{G}_1$. This matter is discussed by Latouche and Ramaswami in [6] and is addressed in our context in Sect. 14.5.

14.4.2 Level-Dependent QBD with Catastrophes

The model in the previous section considered that the file dissemination terminates when no more segments are available for sharing in the system. However, in reality when only an individual segment or an incomplete file remains in the network, no peer is able to retrieve the file completely anymore. Therefore, we now consider that a file is not available for sharing as soon as one of its segments is lost. In this case, the process ends in an absorbing state defined as belonging to $L(0)$ which is defined in this new setting as

$$L(0) = \{(0,0,0),(n,0,0),(0,n,0);n \in \mathbb{N}_0\},$$

where $\mathbb{N}$ (respectively, $\mathbb{N}_0$) is the set of natural numbers (respectively, strictly positive natural numbers). We propose not to differentiate for any $n \in \mathbb{N}_0$ between the states $(n,0,0)$ and $(0,n,0)$, but instead define a kind of metastate labeled $(k,0,0)$ and $(0,k,0)$ that gathers all of these states $(n,0,0)$ and $(0,n,0)$ for $n \in \mathbb{N}_0$, respectively. The subspace $L(0)$ is, thus, composed of three states, that is $\{(0,0,0),(k,0,0),(0,k,0)\}$ and is an absorbing level. Other level state-spaces are for $k \geq 1$:

$$L(k) = \{(i,j,l)\,|\,i,j \in \mathbb{N}, l \in \mathbb{N}_0, i+j+l = k\} \cup \{(i,j,0)\,|\,i,j \in \mathbb{N}_0, i+j = k\}. \tag{14.9}$$

The time to extinction is still equal to the time to absorption and the generator of this new level-dependent QBD is given in (14.10) as follows:

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ A_2^{(1)} & A_1^{(1)} & A_0^{(1)} & 0 & 0 & 0 & \cdots \\ A_3^{(2)} & A_2^{(2)} & A_1^{(2)} & A_0^{(2)} & 0 & 0 & \cdots \\ A_3^{(3)} & 0 & A_2^{(3)} & A_1^{(3)} & A_0^{(3)} & 0 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{14.10}$$

The rates of catastrophe, determined by matrix $A_3^{(k)}$, are given by the transitions and corresponding rates in Table 14.3. Accordingly, matrix $A_2^{(k)}$ becomes as shown in Table 14.4.

The other transitions in matrices $A_0^{(k)}$ and $A_1^{(k)}$ stay the same as previously described in Sect. 14.4.1.1 for the first model, taking care of the states that now belong to the subspace $L(k)$, as defined in (14.9).

Table 14.3 Transitions and rates for matrix $A_3^{(k)}$.

Transitions		Rates
$S_2 > 0:$	$(0,S_2,1) \to (0,k,0)$	d
$S_2 > 0:$	$(1,S_2,0) \to (0,k,0)$	d
$S_1 > 0:$	$(S_1,0,1) \to (k,0,0)$	d
$S_1 > 0:$	$(S_1,1,0) \to (k,0,0)$	d
	$(0,0,1) \to (0,0,0)$	d

Table 14.4 Transitions and rates for matrix $A_2^{(k)}$ with catastrophes.

Transitions		Rates
$S_1 > 1$ or $S_3 > 0:$	$(S_1,S_2,S_3) \to (S_1-1,S_2,S_3)$	$S_1\,d$
$S_2 > 1$ or $S_3 > 0:$	$(S_1,S_2,S_3) \to (S_1,S_2-1,S_3)$	$S_2\,d$
$(S_1 > 0$ and $S_2 > 0)$ or $S_3 > 1:$	$(S_1,S_2,S_3) \to (S_1,S_2,S_3-1)$	$S_3\,d$

The extinction probability can now be computed by extending the results by Bean and Latouche [18] to the level-dependent case. The authors in [18] analyze QBD processes with catastrophes as defined in our setting. However, their phase state-space is of infinite size, whereas in our setting this is no longer the case and makes the problem easier to handle from a numerical viewpoint.

We first define $G_0^{(k)}$ as a matrix whose (i,j)th element is the probability that the process reaches level 0 for the first time in phase j, given that the process starts in phase i of level $k \geq 1$ and levels 1 to $k-1$ are taboo. Let G_k be the matrix whose (i,j)th element is the probability that the process reaches level $k-1$ for the first time in phase j, given that the process starts in phase i of level $k \geq 1$. The extinction probability is then given by G_1 which is here also equal to $G_0^{(1)}$ by definition of this quantity. Moreover, we have for $k \geq 2$ that G_k is given by

$$G_k = \left(A_1^{(k)}\right)^{-1} A_2^{(k)} + \left(A_1^{(k)}\right)^{-1} A_0^{(k)} G_{k+1} G_k. \tag{14.11}$$

Indeed, starting from level k, the QBD may directly move to level $k-1$ with probability $\left(A_1^{(k)}\right)^{-1} A_2^{(k)}$, or it may move up to level $k+1$ with probability $\left(A_1^{(k)}\right)^{-1} A_0^{(k)}$. Upon arrival in level $k+1$, it eventually returns to level k with probability G_{k+1} and then to level $k-1$ with probability G_k. However, the equation for G_1 is slightly different and is given by

$$G_1 = \left(A_1^{(1)}\right)^{-1} A_2^{(1)} + \left(A_1^{(1)}\right)^{-1} A_0^{(1)} \left[G_2 G_1 + G_0^{(2)}\right].$$

If the process moves up to level 2 with probability $\left(A_1^{(1)}\right)^{-1} A_0^{(1)}$ (the second term in this sum), then to reach level 0, it may first return to level 1 with probability G_2 and then move to level 0 with probability G_1. It may also be directly absorbed in level 0 this time without returning to level 1 first. This happens with probability $G_0^{(2)}$. Thus, to compute G_1, we need to know G_2 and $G_0^{(2)}$. More generally, $G_0^{(k)}$ satisfies the following recursive equation:

$$G_0^{(k)} = \left(A_1^{(k)}\right)^{-1} A_3^{(k)} + \left(A_1^{(k)}\right)^{-1} A_0^{(k)} \left[G_{k+1} G_0^{(k)} + G_0^{(k+1)}\right]. \tag{14.12}$$

Its interpretation follows directly from the definition of $G_0^{(k)}$ using the same argument as before. Thus, writing $Q_i^{(k)} = \left(-A_1^{(k)}\right)^{-1} A_i^{(k)}$, $0 \leq i \leq 3$, we have explicitly

$$G_0^{(k)} = \left[I - Q_0^{(k)} G_{k+1}\right]^{-1} \left[Q_3^{(k)} + Q_0^{(k)} G_0^{(k+1)}\right]. \tag{14.13}$$

This implies that to obtain $G_0^{(2)}$ we need $G_0^{(3)}$ and so on. So, we have to truncate the QBD after some level M to be able to start the recursion. We start computing G_M using the logarithmic-reduction algorithm as described in [19]; that is,

$$G_M = \sum_{l=0}^{\infty} \left[\prod_{i=0}^{l-1} U^i_{M-1+2^i} \right] D^l_{M-1+2^l}, \tag{14.14}$$

where the matrices U^l_k and D^l_k are given by (14.5)–(14.8). Accordingly, we obtain the matrices G_{M-1}, G_{M-2}, $\ldots$, G_2 with (14.11). Using (14.13), we finally end up with the following system, which provides us the extinction probability G_1:

$$\begin{aligned}
G_0^{(M)} &= Q_3^{(M)}, \\
G_0^{(M-1)} &= \left[I - Q_0^{(M-1)} G_M \right]^{-1} \left[Q_3^{(M-1)} + Q_0^{(M-1)} G_0^{(M)} \right], \\
&\vdots \\
G_0^{(1)} &= \left[I - Q_0^{(1)} G_2 \right]^{-1} \left[Q_2^{(1)} + Q_0^{(1)} G_0^{(2)} \right] = G_1.
\end{aligned}$$

By truncating the QBD at level M, we actually compute the extinction probability under the taboo of level $M+1$, but a sufficiently large M will provide us a good approximation of this extinction probability.

14.5 Numerical Evaluation

Let us now consider the numerical evaluation of the proposed models, starting with the analysis of the optimistic case. We assume that initially there is a single source sharing both segments in the network, so the system starts at state $(0,0,1)$. The accuracy of our proposed algorithm for computing the extinction probabilities in Sect. 14.4.1 depends on the term l, at which the infinite sum in (14.3) is truncated. Experiments show that in our case the accuracy for $l = 3$ is already sufficient.

The resulting extinction probability as a function over the death rate is illustrated in Fig. 14.3 for file sizes of $F = 9.28$ MB and $F = 6.8$ MB, with $Z = 4.64$ MB as defined earlier being the size of the first segment. The smaller file size has the effect that the second segment is transmitted faster and thus more copies of it exist in the network, which reduces the overall extinction probability slightly. In general, this result can be interpreted as follows. The average death rate d corresponds to the reciprocal of the average sharing time of a peer in the system in seconds. Thus, in order for the content provider to keep a low extinction probability of about 0.01, he should provide incentives that a peer remains in the system for at least 100 s.

We now look at the more pessimistic case that the dissemination stops when at least one segment is no longer available for sharing. In Fig. 14.4, a file size of $F = 9.28$ MB is considered and the death rate d is fixed and equal to 10^{-2}. For the probability that none of both segments are left in the system (i.e., case $(0,0,0)$), we can see that all probabilities are identical and are thus not affected by the truncation level M. However, a slight difference can be seen when we compare the probabilities where only one kind of segment becomes extinct.

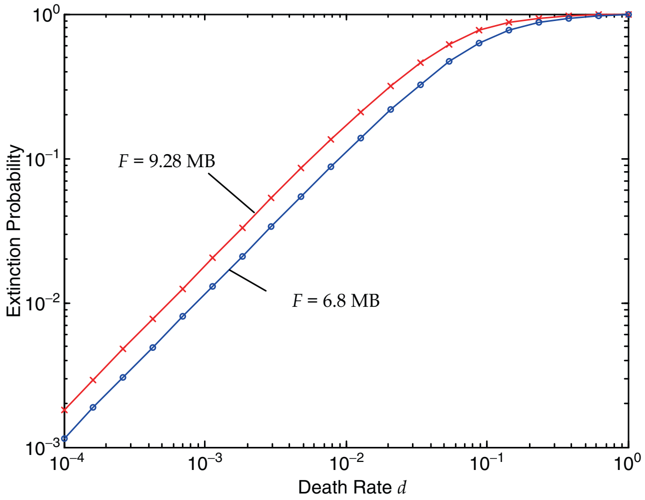

Fig. 14.3 Extinction probability for file sizes $F = 9.28$ MB and $F = 6.8$ MB. When the death rate approaches 1, the extinction probability increases drastically to 1.

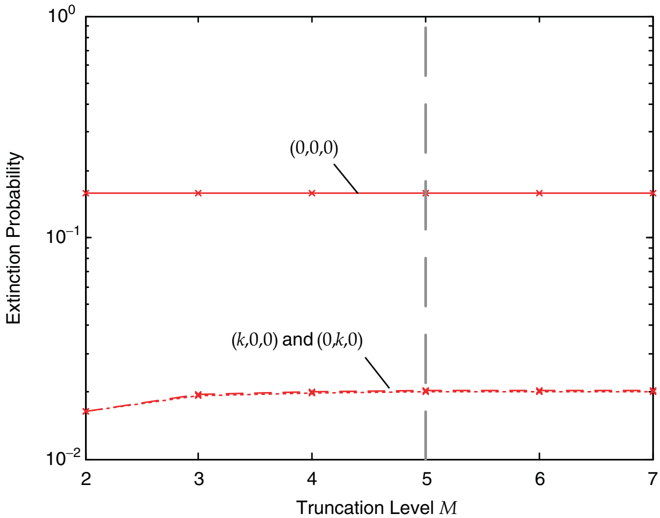

Fig. 14.4 Influence of the truncation level M on the accuracy. A value of about $M = 5$ proves to be accurate enough, so in the following evaluations we use this value as the truncation point.

If we plot the extinction probabilities from the second model with catastrophes over the death rate, we can recognize in Fig. 14.5 that the probabilities to reach $(0,0,0)$ lie above the two curves corresponding to states $(k,0,0)$ and $(0,k,0)$. The reason why they are larger can be interpreted as follows. Initially, the system starts at state $(0,0,1)$, that is, with exactly a single sharing peer. In order to reach the

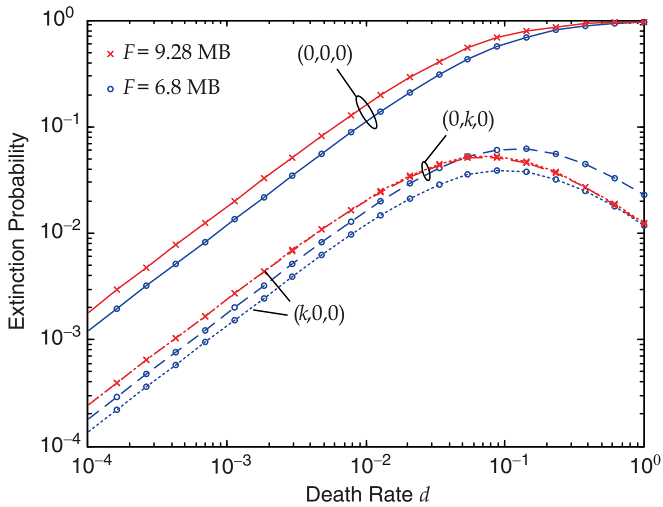

Fig. 14.5 Extinction probabilities with catastrophes for $M = 5$ and file sizes of $F = 9.28$ MB and $F = 6.8$ MB.

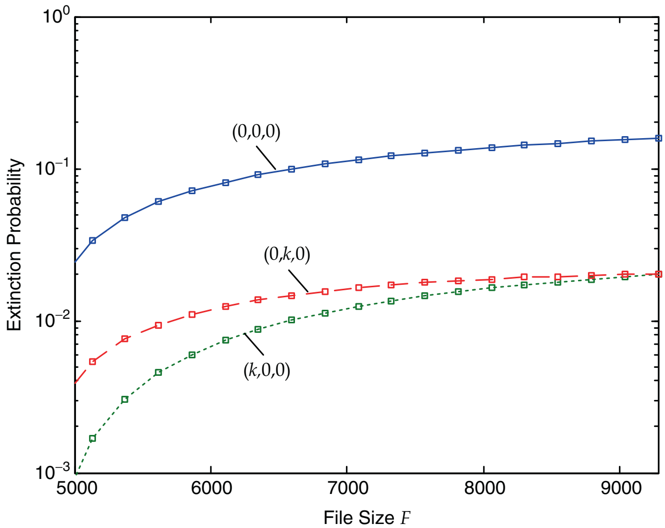

Fig. 14.6 Influence of file size F on the extinction probabilities for $d = 10^{-2}$.

absorbing state $(0,0,0)$, this peer may either make a direct transition by leaving the system or an indirect path by first giving birth to other peers which then all leave after time. On the other hand, in order to reach one of the other absorbing states $(k,0,0)$ or $(0,k,0)$ at least one birth must take place to increment S_1 or S_2, respectively. Thus, a direct transition from $(0,0,1)$ to an absorbing state of that type does not exist in this case, causing a reduction in the weight of the probability.

Additionally, when we look at the shape of the curves, we can recognize that both curves for $(k,0,0)$ and $(0,k,0)$ are identical, when we consider equal segment sizes and the probability for finding and sharing both segments is equal. With $F =$ 6.8 MB the second segment is only half in size of the first, which results in a higher extinction probability of the first segment. The curves lie below the corresponding curves for $F = 9.28$ MB when the death rate d is small. However, in both cases we can see that when the death rate exceeds 10^{-1} the extinction probabilities drop again. At this point it is more likely that the sharing process will stop before any segment is actually downloaded at all; that is

$$d \gg \mu_1(1,1) + \mu_2(1,1),$$

where $\mu_1(1,1) + \mu_2(1,1)$ corresponds to the rate of observing a first new peer with any one of the segments.

The influence of the file size F and, thus, the different size of the second segment is illustrated in Fig. 14.6. We can recognize firstly that for a death rate of $d = 10^{-2}$ the extinction probabilities increase with the file size and, secondly, that when the second segment size is small, the difference between the extinction probabilities of states $(k,0,0)$ and $(0,k,0)$ is large. As expected, when both sizes are equal, both curves approach the same value.

14.6 Conclusions

We provided in this chapter an algorithmically tractable analysis of a level-dependent QBD process with and without catastrophe in terms of the absorption probability, which corresponds to the extinction probability of a file, when we apply the model to file diffusion in unstructured P2P file sharing networks. Numerical results have confirmed that there is a need for the content provider to offer incentives to the peers to encourage sharing and a long sojourn time in the system in order to maintain a sufficiently low extinction probability.

In the future we will use this model to analytically derive further performance measures, especially transient ones such as the distribution of the number of peers present in the system. Furthermore, we would like to enhance the model to consider a more sophisticated peer behavior by including, for example, their willingness to share, impatient peers, and pollution.

References

1. D. Rubenstein and S. Sahu, Can unstructured P2P protocols survive flash crowds?, *IEEE/ACM Transactions on Networking*, vol. 13, no. 3, pp. 501–512, 2005.
2. N. Christin, A. S. Weigend, and J. Chuang, Content availability, pollution and poisoning in file sharing peer-to-peer networks, in *Proc. ACM Conference on Electronic Commerce (EC)*, 2005.

3. A. Binzenhöfer and K. Leibnitz, Estimating churn in structured P2P networks, in *Proc. 20th International Teletraffic Congress (ITC)*, 2007.
4. I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan, Chord: A scalable peer-to-peer lookup service for Internet applications, in *Proc. ACM SIGCOMM*, 2001.
5. S. Hautphenne, K. Leibnitz, and M.-A. Remiche, Extinction probability in peer-to-peer file diffusion, *ACM SIGMETRICS Performance Evaluation Review*, vol. 34, no. 2, pp. 3–4, 2006.
6. G. Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, Philadelphia: ASA-SIAM Series on Statistics and Applied Probability, 1999.
7. S. Saroiu, K. P. Gummadi, R. J. Dunn, S. D. Gribble, and H. M. Levy, An analysis of internet content delivery systems, in *Proc. 5th USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2002.
8. K. Tutschku, A measurement-based traffic profile of the eDonkey filesharing service, in *Proc. 5th Passive & Active Measurement Workshop (PAM)*, 2004.
9. K. Tutschku and H. de Meer, A measurement study on signaling on Gnutella overlay networks, in *Proc. Communication in Distributed Systems (KiVS)*, 2003.
10. M. Izal, G. Urvoy-Keller, E. Biersack, P. Felber, A. A. Hamra, and L. Garces-Erice, Dissecting BitTorrent: Five months in a torrent's lifetime, in *Proc. 5th Passive and Active Measurement Workshop (PAM)*, 2004.
11. T. Hoßfeld, K. Leibnitz, R. Pries, K. Tutschku, P. Tran-Gia, and K. Pawlikowski, Information diffusion in eDonkey-like P2P networks, in *Proc. Australian Telecommunication Networks and Applications Conference (ATNAC)*, 2004.
12. Z. Ge, D. Figueiredo, S. Jaiswal, J. Kurose, and D. Towsley, Modeling peer-peer file sharing systems, in *Proc. 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, pp. 2188–2198, 2003.
13. D. Qiu and R. Srikant, Modeling and performance analysis of BitTorrent-like peer-to-peer networks, in *Proc. ACM SIGCOMM*, 2004.
14. G. Carofiglio, R. Gaeta, M. Garetto, P. Giaccone, E. Leonardi, and M. Sereno, A statistical physics approach for modelling P2P systems, *ACM SIGMETRICS Performance Evaluation Review*, vol. 33, no. 2, pp. 3–5, 2005.
15. X. Yang and G. de Veciana, Service capacity of peer to peer networks, in *Proc. 23rd Conference of the IEEE Communications Society (INFOCOM)*, pp. 2242–2252, 2004.
16. K. Leibnitz, T. Hoßfeld, N. Wakamiya, and M. Murata, On pollution in eDonkey-like peer-to-peer file-sharing networks, in *Proc. 13th GI/ITG Conference on Measurement, Modeling, and Evaluation of Computer and Communication Systems (MMB)*, 2006.
17. K. Leibnitz, T. Hoßfeld, N. Wakamiya, and M. Murata, Peer-to-peer vs. client/server: Reliability and efficiency of a content distribution service, in *Proc. 20th International Teletraffic Congress (ITC)*, 2007.
18. N. Bean and G. Latouche, Numerical approximations for QBDs with infinitely many phases, *Université Libre de Bruxelles, Département d'Informatique, Technical report*, 2007.
19. V. Ramaswami and P. Taylor, Some properties of the rate operators in level dependent quasi-birth-and-death processes with a countable number of phases, *Communications in Statistics: Stochastic Models*, vol. 12, no. 1, pp. 143–164, 1996.