

Modeling and Analysis of Upload Utilization in BitTorrent System

J.W. Zhang^{1, a}, Y.C. Wang^{2, b}, W. Xing^{*, 1, c} and D.M. Lu^{1, d}

¹College of Computer Science and Technology, Zhejiang University, China

²Library and Information Center, Zhejiang University, China

{ ^ajanyway, ^bychwang, ^cwxing, ^dldm } @zju.edu.cn

*Corresponding author

Keywords: BitTorrent, Piece diffusion, Upload utilization, TFT

Abstract. Upload utilization is one of the most significant performance metrics in BitTorrent system, whereas there has been no accurate model which is suitable for it yet. For this reason, the delicate relation between some key parameters (maximum concurrent uploads, upload bandwidth, TFT round etc.) in BitTorrent is investigated, furthermore the piece diffusion model and the upload utilization model are proposed. Based on these models, we explore the effect of TFT round for the first time and find that the magnitude of TFT round has a significant influence on the transient stage as well as the end stage. Additionally, we show that when the seed's upload bandwidth is fully utilized, to achieve optimal performance, each peer's upload bandwidth should be proportional to its maximum concurrent uploads, and the file size should be large enough compared with the swarm size.

Introduction

BitTorrent is currently the most effective P2P protocol for scalable content distribution, its high efficiency mainly derives from the local rarest first (LRF) policy for piece selection and the tit-for-tat (TFT) policy for peer selection [1]. The high complexity and randomness of BitTorrent rests above all in its large quantity of parameters, e.g., piece size, maximum concurrent uploads, upload bandwidth, TFT round, number of peers, file size, etc. To achieve optimal performance, lots of studies have concentrated on modeling, analysis and improvement for BitTorrent system.

Qiu et al. [2] present a fluid-flow model in BitTorrent which reveals the dynamic balance between peers' arrival and peers' departure. However, they only focus on scalability and incentive mechanism in the steady stage. Yang et al. [3] propose a model based on age dependent branching processes, and argue that the transient stage in P2P system exhibits exponential growth in service capacity. But they make an unrealistic assumption that the source peer has the same upload bandwidth as other peers. Defigueiredo et al. [4] demonstrate that the TFT strategy in P2P system scales at a linear rate rather than an exponential rate. But we find this would not apply to the TFT policy employed in BitTorrent. In order to improve the fairness of BitTorrent, Thommes et al. [5] propose a modification in which the number of connections a peer maintains varies with its upload bandwidth instead of a fixed number. But it is not amenable to theoretical analysis. Marciniak et al. [6] show that piece size is critical for performance, as it determines the degree of parallelism in the system.

In previous works which focused on modeling, analysis, and improvement for BitTorrent, there has been no accurate model suitable for the upload utilization, especially when the seed's upload bandwidth is fully utilized, and the delicate relation between some key parameters has not been revealed either. In this paper, we present the piece diffusion model and the upload utilization model. By using the models, we explore the effect of TFT round on system performance for the first time and show the delicate relation between some key parameters in order to achieve optimal performance. All these models and derived results can be used to further understand the properties of BitTorrent and to optimize the performance in BitTorrent-like systems.

Model

Assumptions and Scenario. In this paper, as we mainly concentrate on the piece diffusion process and the upload utilization of the system, the incentive mechanism will not be related. In order to conveniently research and without loss of generality, we modify the seeding strategy in two aspects, first, we make the seed contribute evenly to a fixed number of leechers in every TFT round, this is in accordance with the time-based seeding strategy [7] employed by many BitTorrent clients. Second, we make the seed always send the piece that it has sent the least to the swarm, which is first proposed in [8]. Therefore, before the first file copy has been sent, the seed is always sending various pieces to its neighbors, i.e. each piece will be sent only once during a single round.

In previous studies [3, 4] which focused on the service capacity and the performance improvement, almost all researchers make an assumption that the seed’s capacity is no less than the leechers’, i.e. leechers’ upload bandwidth is the performance bottleneck. In this case, all leechers’ upload bandwidth will be fully utilized. As a complement and expansion to previous works, our study is based on the condition that leechers’ upload bandwidth is no less than the seed’s, i.e. $U \geq C$, which implies that the seed’s upload bandwidth will be fully utilized.

We assume the swarm size to be N (i.e. the number of peers), it consists of an initial seed with an upload bandwidth C and leechers with identical upload bandwidth U . All peers’ download bandwidth is unlimited. The file size is Z , divided into M pieces, hence the piece size p is $\frac{Z}{M}$. The maximum concurrent uploads of the seed and the leechers are n and v respectively. The TFT round is denoted as R (10 seconds by default).

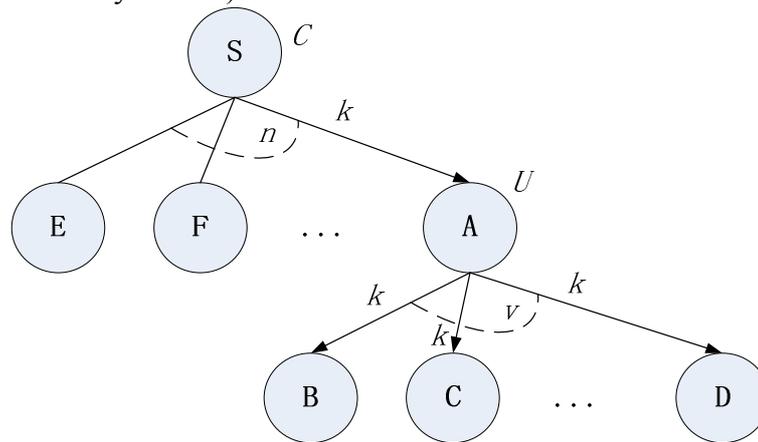


Fig.1 Piece Diffusion Process.

Maximum Concurrent Uploads. In Eq. 1, k is the number of pieces that the seed can send to a single leecher in one TFT round.

$$k = \frac{CR}{pn} \tag{1}$$

As to peer A in Fig.1, in order to distribute the newly-received pieces among its neighbors in the following TFT round, first, peer A (see Fig. 1) should receive at least one complete piece from the seed (peer S) in the last round, i.e. $k \geq 1$, and second, the upload bandwidth of peer A should be at least $\frac{vpk}{R}$ so that it can be able to send kv pieces to its v selected neighbors in the next TFT round, thus we get Eq. 2.

$$U = \frac{vpk}{R} \tag{2}$$

Substituting k into Eq.2, then we have Eq.3.

$$\frac{C}{n} = \frac{U}{v} \tag{3}$$

Eq.3 indicates that, to make full use of the total upload bandwidth in the system and to achieve optimal performance, each peer's upload bandwidth should be proportional to its maximum concurrent uploads.

File Size and Swarm Size. The time a piece is diffused to all peers is denoted by t_p (based on the unit of round) and the time the first file copy is sent into the swarm by the seed is denoted by t_f (based on the unit of round).

According to our modified seeding strategy, If we assume that the seed always utilizes its full upload bandwidth, then t_f can be calculated as Eq. 4.

$$t_f = \frac{Z}{CR} - 1. \tag{4}$$

In Fig.1, a new piece p_i will be replicated only once by the seed S as the seed always sends the piece that it has sent the least. When peer A get p_i , as the pieces sent before p_i have already spread all over its neighbors, p_i must be the rarest piece to upload. If the leecher's upload bandwidth is large enough (see Eq. 2), it will distribute its newly-received pieces to all its selected v neighbors in one round's time. Thus we have Eq. 5.

$$t_p = \lceil \log_v(N-1) \rceil + 1. \tag{5}$$

Obviously, we can find that t_p and t_f are determined by the swarm size N and the file size Z respectively. If $t_f < t_p$, i.e., before the pieces entered the swarm first diffuse among all peers, the first file copy has been already sent by the seed. Subsequently, the upload utilization has already begun to decrease without reaching its maximum value. Therefore, in practice, in order to reach the maximum upload utilization, we had better make sure t_f is greater than t_p , i.e. Eq. 6, which implies that the file size should be large enough compared with the swarm size.

$$t_f > t_p. \tag{6}$$

Piece Diffusion Model. In Fig.2, each curve represents the diffusion processes of kn pieces sent by the seed during one round, and these processes are fully overlapped on our assumption of ideal condition. To facilitate discussion, the pieces sent in round i ($i \in [0, t_f]$) are regarded as the piece set i . The time spans of $[0, t_p]$, $[t_p, t_f]$ and $[t_f, t_e]$ are called the transient stage, the steady stage and the end stage respectively. We denote the replicas of all pieces in piece set i as $rp_i(t)$. If the leechers' upload bandwidth is larger than in Eq.2, the seed's upload bandwidth can be fully utilized, then $rp_i(t)$ can be calculated as Eq. 7.

$$rp_i(t) = \begin{cases} 1, & 0 \leq t \leq i \\ v^{t-i-1} + 1, & i + 1 \leq t \leq i + t_p \\ N, & t > i + t_p \end{cases} \tag{7}$$

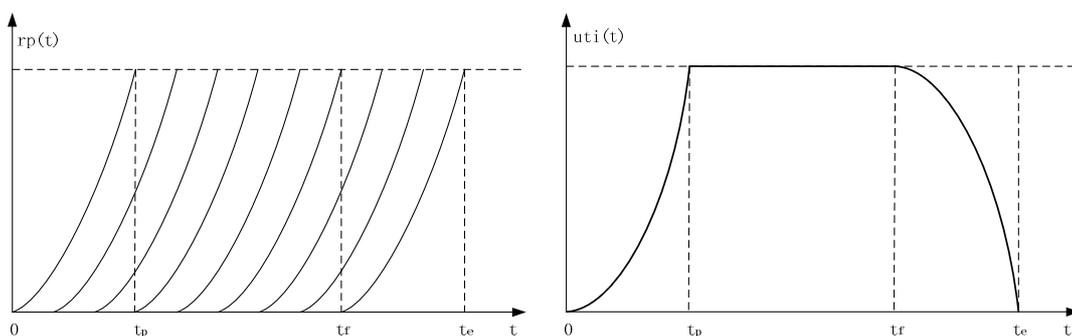


Fig.2 Piece Diffusion Model and Upload Utilization Model.

Upload Utilization Model. Upload utilization is one of the most significant performance metrics in BitTorrent system. The upload utilization in round t is defined as the ratio of total data transmit rate to total upload bandwidth of all peers. As the total data transmit rate can be expressed as the ratio of total data amount transmitted during last round (i.e. from $t-1$ to t) to the TFT round R , let U_t be the total upload bandwidth of all peers, i.e. $U_t = (N-1)U + C$, then we can calculate $uti(t)$ as Eq. 8.

$$uti(t) = \frac{knP \sum_0^{t_f} [rp_i(t) - rp_i(t-1)]}{U_t R}. \quad (8)$$

With the assumption that the diffusion process of each piece is exactly the same, the replicas of piece set $i+1$ at round $t+1$ is equal to the replicas of piece set i at round t , i.e. Eq.9 holds.

$$rp_i(t) = rp_{i+1}(t+1). \quad (9)$$

Combining Eq. 9 with Eq.8, we can calculate the total data amount transmitted from $t-1$ to t as Eq. 10.

$$knP \sum_0^{t_f} [rp_i(t) - rp_i(t-1)] = \begin{cases} knP(rp_0(t)-1), & 0 \leq t \leq i \\ knP(N-1), & i+1 \leq t \leq i+t_p \\ knP(N-rp_f(t)), & t > i+t_p \end{cases} \quad (10)$$

Substituting Eq.10 into Eq.8, we can then calculate the upload utilization function $uti(t)$ as Eq.11.

$$uti(t) = \begin{cases} \frac{C}{U_t} v^{t-1}, & t \in [0, t_p] \\ \frac{C}{U_t} (N-1) \approx \frac{C}{U}, & t \in [t_p, t_f] \\ \frac{C}{U_t} (N-1-v^{t-t_f}), & t \in [t_f, t_e] \end{cases} \quad (11)$$

Eq.11 implies that leechers' upload bandwidth will not be fully utilized when it is larger than the seed's in the steady stage, it only effects on the transient stage and the end stage. Unlike the discussion about TFT strategy in [4], which demonstrates the TFT strategy cannot scale as well as a fully cooperative strategy, Eq.11 shows that the TFT strategy employed in BitTorrent has indeed an exponential increasing rate in the transient stage whereas an exponential decreasing rate in the end stage.

As t is based on the unit of R , we can find that the time span of transient stage decreases with the decreasing of R , i.e., the smaller R is, the higher the increasing rate of $uti(t)$ is. According to Eq.4, the time span from 0 to t_f is not affected by R . In Fig.2 we have $t_e = t_f + t_p$, so the time span of the end stage (from t_f to t_e) is as long as the transient stage. Similarly, the smaller R is, the higher the decreasing rate of $uti(t)$ is in the end stage.

Simulation and Analysis

We employ a discrete event-driven simulator BitSim [9]. All parameters are set to the default values unless otherwise specified. The swarm size is 41, all peers (including the seed) have a 256 KB/s upload bandwidth and an unlimited download bandwidth. All peers join the swarm in a flash-crowd way and depart from the swarm as soon as they finish their download.

Effect of the TFT Round. We find that the TFT round R is as critical as the piece size, it will also determine the degree of parallelism in the system. We define the average upload utilization (AUU) as the metric of system performance, which can be calculated as an average of $uti(t)$ from 0 to t_e .

We let the seed and the leecher have the same maximum concurrent uploads n , then we investigate the change of AUU on R under different piece size p and different n . Fig.3 shows that, when v is 4 and the piece size is 256K, the AUU decreases with the increasing of R . The major reason for this is as discussed above, i.e. a smaller R makes both the transient stage and the end stage shorter whereas the steady stage, which has a high upload utilization, is prolonged. Particularly, the AUU drops sharply when R is 2. Using Eq. 1, we can see that the seed has to send 2 pieces to 4 leechers in one round. As the seed equally allocates its upload bandwidth to 4 neighbors, each leecher will not get a complete piece during one round. This will severely hinder the leechers from sharing pieces with other leechers as soon as possible and will reduce the AUU seriously.

As a result, we can conclude that, when decreasing the value of R to achieve better performance, we should make k be larger than 1. According to Eq.1, there is a lower bound on R , i.e. $R \geq \frac{np}{C}$. This makes peers get complete pieces more quickly so that they can share the newly-recieved pieces with others as soon as possible.

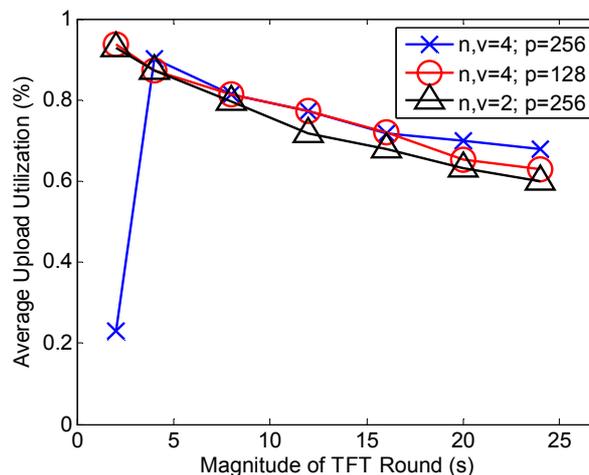


Fig.3 Average Upload Utilization vs. Magnitude of TFT Round.

Model Validation. Fig. 4 shows the piece diffusion process and the upload utilization curve when R is 4. We can find that both the piece diffusion model and the upload utilization model fit the simulation results very well, the diffusion processes of pieces in the same piece set have a high degree of overlap, and the diffusion processes of pieces in different piece set are closely parallel. Notice that we only plot the diffusion process of pieces in piece set 0 (i.e. $rp_0(t)$) for clarity. As a peer cannot always obtain required pieces from its neighbors, the maximum upload utilization in simulation is a little lower than the optimal value in the model.

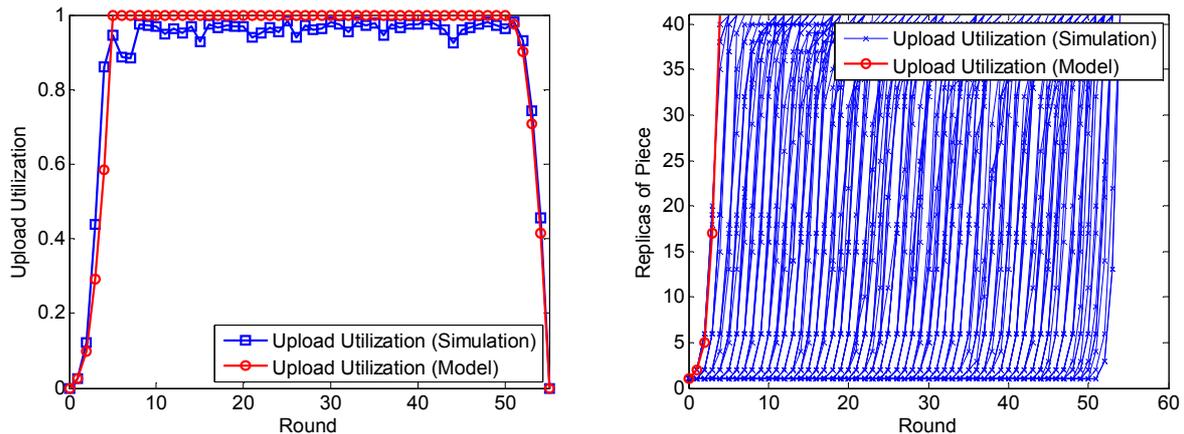


Fig. 4 Replicas of Piece and Upload Utilization.

Summary

In this paper, we propose the piece diffusion model and the upload utilization model in BitTorrent system. Based on these models, we explore the effect of the TFT round and show that, the magnitude of TFT round has a significant influence on both the transient stage and the end stage. Additionally, we find that, to achieve optimal performance, each peer's upload bandwidth should be proportional to its maximum concurrent uploads, and the file size should be large enough compared with the swarm size.

This paper is supported by the National Key Technology R&D Program (No. 2007BAH11B06) and the Next Generation Internet special of NDRC (No. CNGI2008-112).

References

- [1] B. Cohen, Incentives build robustness in BitTorrent, Proc. of the 1st Workshop on Economics of Peer-to-Peer Systems. Berkeley, 2003.
- [2] D. Qiu and R. Srikant, Modeling and Performance Analysis of BitTorrent-Like Peer-to-Peer Networks, Proc. of SIGCOMM'04, New York, 2004, pp. 367–378.
- [3] X. Yang and G. De Veciana, Performance of peer-to-peer networks: Service capacity and role of resource sharing policies, Performance Evaluation, vol. 63, 2006, pp. 175-194.
- [4] D. Defigueiredo, B. Venkatachalam and S. F. Wu, Bounds on the performance of P2P networks using tit-for-tat strategies, Proc. of P2P'07, Galway, 2007, pp. 11-18.
- [5] R. Thommes and M. Coates, BitTorrent fairness: analysis and improvements, Proc. of 4th Workshop Internet, Telecom. and Signal, Noosa Heads, 2005.
- [6] P. Marciniak, N. Liogkas, A. Legout and E. Kohler, Small is not always beautiful, Proc. IPTPS'08, Berkeley, 2008.
- [7] X. Chen and S. A. Jarvis, Analysing bittorrent's seeding strategies, Proc. of 12th IEEE International Conference on Computational Science and Engineering, Vancouver, 2009, pp. 140-149.
- [8] A. Bharambe, C. Herley, and V. Padmanabhan, Understanding and deconstructing BitTorrent performance, Microsoft Corp., Redmond, WA, Tech. Rep. MSR-TR-2005-03, 2005.
- [9] Information on <http://planete.inria.fr/software/BitSim/>