Submitted version

Peer-to-Peer Networking and Applications manuscript No.
(will be inserted by the editor)

Optimality Analysis of Locality-Aware Tit-for-Tat-Based P2P File Distribution

Yohei Nishi · Masahiro Sasabe · Shoji Kasahara

Received: date / Accepted: date

Abstract In the current Internet, periodic software update is essential to maintain systems secure from malicious attacks. In case of proliferated software, e.g., Operating System (OS), the distribution server for update tends to be a bottleneck, due to the concentration of users' requests. To tackle this problem, several systems, e.g., Windows Update, have been applying the Peer-to-Peer (P2P) file distribution technique where clients called peers upload retrieved fragments of the original file, i.e., pieces, to others. However, peers may not be willing to upload pieces to others, due to their own communication overhead. BitTorrent has adopted the Tit-for-Tat (TFT) strategy in game theory, which encourages peers to exchange an equivalent number of pieces among each pair of peers. In recent years, the optimality of TFT-based P2P content distribution, i.e., file distribution and streaming, has been analyzed with the help of Integer Linear Programming (ILP). In this paper, considering the fact that the communication overhead inside a group, e.g., LAN or Autonomous System (AS), is much less than that of between groups, we model locality-aware TFT-based P2P file distribution where the TFT constraint is relaxed for intra-group communications. We further formulate the optimal piece flow determination problem as ILP in the similar way to the existing work. Through numerical results, we show that the locality-aware TFT-based P2P file distribution can achieve quasi-optimal average file download time when the upload capacity of the server is a bottleneck and the number of groups is moderate.

This paper is an expanded version of the paper, which will be presented at IEEE Consumer Communications & Networking Conference (CCNC 2020) [20].

Yohei Nishi (Corresponding author), Masahiro Sasabe, Shoji Kasahara
Graduate School of Science and Technology, Nara Institute of Science and Technology
8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan
Tel.: +81-743-5361
Fax: +81-743-5369
E-mail: nishi.yohei.no0@is.naist.jp, {m-sasabe, kasahara}@ieee.org

Keywords P2P file distribution · Tit-for-Tat strategy · Locality-awareness · Optimal piece flow · Integer Liner Programming (ILP)

1 Introduction

Nowadays, periodic software update is essential to maintain the system secure against malicious attacks. When the software is used by many users (e.g., Operating System (OS)) and/or new update files (e.g., security patch files) are released, users' accesses concentrate on the distribution server, which causes flash crowd [6] and makes the server a bottleneck in terms of CPU and network resources. To tackle the scalability issue, several systems (e.g., Windows update) have been applying the Peer-to-Peer (P2P) technology to assist the distribution by end-user devices called peers [2]. In the P2P file distribution, a file is divided into multiple fragments called pieces. Peers can download pieces from the server as well as from other peers. If the peers cooperatively exchange pieces with each other, the server can achieve the file distribution that can adapt to the dynamical change of user demands without investments in equipment.

However, peers are user devices and thus their behavior depends on the corresponding users' decision making. In particular, sending pieces to others consumes their own upload capacity, which may not only degrade the quality of other network services running on their devices but also may require additional monetary cost under the access fee plan depending on the usage. As a result, it has been pointed out that many peers tend to become free riders that hesitate to upload pieces to others [4]. Since the increase of free riders decreases the total upload capacity of the system, incentive mechanisms are essential to promote the mutual cooperation among peers.

To alleviate the free-riding problem, several P2P file distribution systems, e.g., BitTorrent [1], applies the Tit-for-Tat (TFT) strategy in game theory [16]. Simply speaking, the TFT strategy in the P2P file distribution can be regarded as the equivalent piece exchange where each peer has to provide pieces with others so as to retrieve its demanding pieces from them. Although the TFT constraint encourages the mutual cooperation among peers, it may also degrade the distribution performance, due to the strict equivalent piece exchanges between peers.

In recent years, the optimality of the TFT-based P2P content distribution and its optimal piece flow have been analyzed [13, 22, 23]. The piece flow is the information that describes the transfer timing of each piece from a node (i.e., server or peer) to a peer. Hasegawa et al. first formulated the piece-flow decision problem to minimize the average file download time as Integer Linear Programming (ILP) [13]. In [22], Sasabe found that the piece-flow decision problem for minimizing the average play-out delay in the TFT-based P2P streaming could also be modeled as ILP. The impact of the P2P (logical) topology on the optimality of the TFT-based P2P content distribution was also analyzed [23]. These existing work obtained the optimal solutions by solving their problems using the existing solver, i.e., CPLEX [3], and revealed the optimal behavior (strategy) of the server and peers, which yields the optimal piece flow by analyzing the optimal solutions.

However, these existing work did not consider the physical proximity among peers (e.g., LAN and Autonomous System (AS)), which would affect the cooperative relationship among peers. In the following, we refer to a set of physically close peers as a *group*. As mentioned above, the free-riding behavior tends to come from avoiding the bandwidth consumption caused by uploading pieces. In case of the intra-group communication, each peer can easily send pieces to other group members with abundant network bandwidth, and thus the cost of uploading is negligible. On the other hand, in case of the inter-group communication, peers potentially tend to be free riders, due to the limited and non-free upload capacity of their access links, as in the existing work.

Moreover, the traffic between peers that belong to different ASes may impose the transit cost on the related Internet Service Providers (ISPs). In general, the transit cost is imposed on the lower tier ISP when it uses the transit link to the upper tier ISP and it is determined according to the q-percentile charging policy [12]. In the q-percentile charging policy, the traffic within a fixed period (e.g., 5 minutes) is periodically measured and sorted in ascending order. The transit cost is determined based on the $q\%$-th largest amount of traffic. Thus, it is important to reduce the peak traffic so as to suppress the transit cost. From the viewpoint of the P2P content distribution, the reduction of the transit cost will contribute to the ISP-friendly content distribution.

In this paper, we focus on the locality-aware TFT-based P2P file distribution where the TFT constraint is relaxed for intra-group communication. After modeling it as a discrete-time system, we formulate the optimal piece flow determination problem as three-step ILP by taking account of the three different players in the P2P file distribution, i.e., server, peers, and ISPs, in this order. We first aim to minimize the average file download time among peers from the viewpoint of the server. Next, we further try to minimize the average inter-group traffic among peers while achieving the first optimization, which will contribute to the reduction of peers' costs. Finally, we attempt to minimize the peak traffic between groups while satisfying the first and second optimization, from the viewpoint of ISPs. Through numerical results obtained by solving the ILP, we reveal the fundamental characteristics of the locality-aware TFT-based P2P file distribution.

The main contributions of this paper are as follows:

1. We develop a discrete-time model of the locality-aware TFT-based P2P file distribution as a relaxed TFT model. We further formulate three-step ILP to determine the optimal piece flow where the average file download time, the average inter-group traffic, and the peak traffic between groups are minimized in this order. We obtain the optimal piece flow by solving the ILP using the existing solver.
2. Through numerical results, we reveal the following: (1) the relaxed TFT model can achieve the average file download time competitive with no TFT

model, with the help of altruistic intra-group piece distribution of peers, (2) the upload capacity of the server has much larger impact on the average file download time than that of peers but a relatively small server's upload capacity can drastically shorten the average file download time, (3) the relaxed TFT model with a moderate number of groups can achieve better average file download time than the existing strict TFT model, and (4) both the average and peak inter-group traffic of the relaxed TFT model tend to be larger than those of no TFT model, due to the TFT constraint on inter-group communication.

The rest of the paper is organized as follows. Section 2 gives related work. In Section 3, we model the locality-aware TFT-based P2P file distribution and formulate the optimal piece flow determination problem as three-step ILP. After demonstrating numerical results in Section 4, we give conclusions in Section 5.

2 Related work

2.1 Incentive mechanisms in P2P content distribution

In P2P content distribution, to alleviate the free riding problems, incentive mechanisms play an important role to encourage peers to cooperate with each other. BitTorrent first adopted the TFT strategy in game theory to achieve mutual cooperation between each pair of peers. Wu et al. designed an incentive mechanism for P2P file sharing based on social network and game theory, where a server gives counters to peers by considering their resource contribution [27]. In [17], the authors designed a credit-based incentive mechanism for heterogeneous P2P networks, where a Stackelberg game is considered to model the optimal pricing and purchasing strategies. Lu et al. proposed a dynamic reward-based mechanism where the reward for each peer decreases (resp. increases) with increase (resp. decrease) of cooperative peers [18]. Ghaderzadeh et al. proposed a method to detect free riders where peers dynamically change their neighbors according to the knowledge about them [11]. In [15], the authors proposed a repeated game model for P2P live streaming where peers are coerced to cooperate by a Striker strategy. In this paper, we focus on the TFT strategy as the incentive mechanism because it can easily be realized only based on the observation at each peer and is robust against cheating behavior.

2.2 Performance evaluation of TFT-based P2P content distribution

There are many studies on performance evaluation of the BitTorrent(-like) systems [28]. In [6], the authors formulated an analytical model to estimate the available bandwidth of the P2P file distribution systems under flash crowd. In [4], the authors proposed a game theory model to mitigate the impact of free riders on the BitTorrent performance. Ferragut et al. proposed partial differential equation to derive the system size and download time of dynamic P2P file sharing network [10]. Silva et al. proposed models to calculate the system throughput of the P2P swarming systems under distinct policies for both peer and piece selection [25].

In recent years, the optimality of the TFT-based P2P content distribution has been studied with the help of ILP [13, 22, 23]. In [13], the authors clarified the optimality of the rarest-first strategy used in BitTorrent through the analysis of the optimal solution in TFT-based P2P file distribution. In addition, it was also revealed that the server should distribute pieces to peers with lower upload capacity at the early stage of distribution, so as to minimize the average file download time. In [22], the optimality of TFT-based P2P streaming was also analyzed . The author revealed that the server should distribute pieces to peers in the descending order of peer's upload capacity at early stage of streaming to minimize the average play-out delay. The impact of the logical network topology among peers on the system performance was also investigated in [23].

In this paper, we focus on how the TFT relaxation based on the physical proximity among peers will affect the system performance of the TFT-based P2P file distribution.

2.3 Locality-aware P2P content distribution

Since a P2P network is constructed over the top of the underlying physical network, e.g., the Internet, the transmission rate and latency between two peers highly depend on their physical closeness. If neighboring peers over the P2P network are physically close, they tend to be able to enjoy high transmission rate with low latency as well as the transit cost between ISPs will also be reduced. Thus, traffic localization on P2P content distribution plays an important role to not only improve the communication quality among peers but also reduce costs for both peers and ISPs.

In recent year, Microsoft has been starting to apply the locality-aware P2P technologies to assist Windows update, in addition to the conventional client–server delivery scheme [2]. Specifically, in the locality-aware P2P delivery scheme, user PCs can send and receive update files with other PCs in the same sub network, e.g., LAN.

In the research field, locality-aware P2P systems have also been studied [5, 8, 9, 14, 19, 24, 26, 29, 30]. Huang proposed the peer selection strategy based on the multi objective optimization, which considers both the system performance and traffic localization [14]. Magharei et al. proposed a scheduling scheme for P2P streaming to reduce inter-ISP traffic [19]. Costa et al. proposed a locality-aware BitTorrent protocol for enterprise networks composed of multiple LANs [9]. Blond et al. analyzed the impact of the number of connections between peers belonging to different ISPs on the download time and communication overhead [5]. Seibert et al. analyzed the localization of P2P systems from the viewpoint of ISP's benefits [24]. Cuevas et al. analyzed

Submitted version

the characteristics of localized P2P systems through both the mathematical model and actual data set collected by BitTorrent [8]. In [30], the authors investigated the relationship between the bandwidth consumption of the servers and the inter-ISP traffic in the hybrid P2P-Cloud streaming, with the help of the stochastic optimization. Zhang et al. evaluated the P2P content distribution assisted by the ISP information from the view points of intra-ISP traffic, economic benefits, and user efficiency [29]. Wang et al. measured and analyzed the impact of locality on P2P content distribution through the simulations on the PlanetLab [26].

In this paper, we focus on the fact that the traffic cost between physically close nodes tend to be low, and analyze the performance of the locality-aware TFT-based P2P file distribution where the TFT constraint among physically close peers is relaxed.

3 Modeling locality-aware TFT-based P2P file distribution and ILP formulation of determination of its optimal piece flow

In this section, we model the locality-aware TFT-based P2P file distribution and formulate the determination of its optimal piece flow as third-step ILP, which tries to optimize the performance of server, peers, and ISPs, in this order.

3.1 Model

As in [13], the TFT-based P2P file distribution is modeled as a discrete-time system. In the system, there are a server, denoted by $\mathcal{N}_{\mathrm{D}} = \{0\}$, and N_P peers, denoted by $\mathcal{N}_{\mathrm{P}} = \{1, \ldots, N_{\mathrm{P}}\}$. In what follow, the server and peers are collectively called *nodes*. For simplicity, we define the set of nodes as $\mathcal{N} = \mathcal{N}_{\mathrm{D}} \cup \mathcal{N}_{\mathrm{P}}$ whose size is $N = |\mathcal{N}|$. Peers play two kinds of roles, *leechers* and *seeds*, depending on their file retrieval states. Leechers are peers under file downloading while seeds are those that finish it. In actual systems, many seeds tend to stay in the system during a relatively short time [21]. In this paper, we assume that each peer leaves the system just after becoming a seed. The content file consists of M fragments, i.e., pieces, denoted by $\mathcal{M} = \{1, \ldots, M\}$. Without loss of generality, we consider the piece size is identical and set to be one.

The two sets of time steps are defined as $\mathcal{T} = \{0, \ldots, T\}$ and $\mathcal{T}^{+} = \{1, \ldots, T\}$, respectively. T should be large enough to ensure that all peers complete file downloading. The derivation of T is discussed in [13].

Next, we focus on the physical proximity (i.e., group relationship) between nodes. Suppose that there are G groups. Each node belongs to an exact one group and the group relationship between node i and node j is given as input parameters:

$$G_{i,j} = \begin{cases} 1, & \text{if } i \text{ and } j \text{ belong to the same group,} \\ 0, & \text{otherwise.} \end{cases}$$

Note that the server $i \in \mathcal{N}_{\mathrm{D}}$ is separated from peers, i.e., $G_{i,j} = 0$ ($j \in \mathcal{N}_{\mathrm{P}}$). From the definition, $G_{i,j} = G_{j,i}$ ($i, j \in \mathcal{N}$). Each node $i \in \mathcal{N}$ can upload at most C_i^{out} pieces per unit time to peers in different groups, i.e., $\{j \in \mathcal{N}_{\mathrm{P}} \mid G_{i,j} = 0\}$. For simplicity, we assume that the inter-group upload capacity of node i, C_i^{out}, is a natural and finite number. Considering the asymmetry between uplink and downlink channel speeds in the current Internet (e.g., ADSL and cable Internet), we assume that the inter-group upload capacity of each peer becomes the bottleneck compared with its inter-group download capacity. Since the intra-group communication (e.g., LAN) is much faster than the inter-group communication, we assume that the intra-group upload capacity of peer $i \in \mathcal{N}_{\mathrm{P}}$, denoted by C_i^{in}, is much larger than C_i^{out}.

The initial state of the system at $t = 0$ can be represented by the piece possession of each node, $z_{0,k,i}$ ($k \in \mathcal{M}, i \in \mathcal{N}$). For example, if the server only has the whole content, $z_{0,k,i} = 1$ ($k \in \mathcal{M}, i \in \mathcal{N}_{\mathrm{D}}$) and $z_{0,k,i} = 0$ ($k \in \mathcal{M}, i \in \mathcal{N}_{\mathrm{P}}$) should hold.In the succeeding time steps $t \in T^+$, the piece flow among nodes determines the system dynamics, which is given by the decision variables $x_{t,k,i,j}$ ($t \in \mathcal{T}, k \in \mathcal{M}, i, j \in \mathcal{N}$):

$$x_{t,k,i,j} = \begin{cases} 1, & \text{if node } i \text{ sends piece } k \text{ to node } j \text{ at time } t, \\ 0, & \text{otherwise.} \end{cases}$$

With the initial condition $z_{0,k,i}$ and the decision variables $x_{t,k,i,j}$, we can also derive two kinds of variables, $z_{t,k,i}$ ($t \in \mathcal{T}^+, k \in \mathcal{M}, i \in \mathcal{N}$) and $y_{t,i}$ ($t \in \mathcal{T}, i \in \mathcal{N}$). $z_{t,k,i}$ denotes the piece possession of node i:

$$z_{t,k,i} = \begin{cases} z_{t-1,k,i} + \sum_{j \in \mathcal{N}} x_{t,k,j,i}, & \text{if } i \in \mathcal{N}_{\mathrm{P}}, \\ 1, & \text{otherwise.} \end{cases}$$

The variables $z_{t,k,i}$ represents that the server $i \in \mathcal{N}_{\mathrm{D}}$ always has all pieces. In the case of peers, $z_{t,k,i} = 1$ if peer i has piece k at time t, and otherwise $z_{t,k,i} = 0$.

The variables $y_{t,i}$ denotes the role of node i:

$$y_{t,i} = \begin{cases} 1 - \prod_{k \in \mathcal{M}} z_{t,k,i}, & \text{if } i \in \mathcal{N}_{\mathrm{P}}, \\ 0, & \text{otherwise.} \end{cases}$$

In case of the server $i \in \mathcal{N}_{\mathrm{D}}$, $y_{t,i} = 0$ ($t \in \mathcal{T}$). On the other hand, if peer $i \in \mathcal{N}_{\mathrm{P}}$ is a leecher at time t ($t \in \mathcal{T}$), $y_{t,i} = 1$, and otherwise, $y_{t,i} = 0$. The notations used in the model are summarized in Table 1.

Table 1 Notations in the model.

Notation	Definition
$\mathcal{N}_{\mathrm{D}}$	The set of server,$\{0\}$
$\mathcal{N}_{\mathrm{P}}$	The set of peers,$\{1, 2, \ldots, N_{\mathrm{P}}\}$
$\mathcal{N}$	The set of nodes, $\mathcal{N} = \mathcal{N}_{\mathrm{D}} \cup \mathcal{N}_{\mathrm{P}}$, $(N = \lvert\mathcal{N}\rvert)$
$\mathcal{M}$	The set of pieces, $\{1, 2, \ldots, M\}$
$G_{i,j}$	Group relationship between peer i and peer j (1: they are in the same group, 0: otherwise)
C_i^{in}	Upload capacity of peer i to group members
C_i^{out}	Upload capacity of peer i to non-group members
$x_{t,k,i,j}$	Decision variables of piece flow
$z_{t,k,i}$	Variables of piece possession (1: possession, 0: missing)
$y_{t,i}$	Variables of peer's role (1: leechers, 0: server/seeds)
τ_i	File download time of peer i
$\overline{\eta}$	Average inter-group traffic among peers
η_{peak}	Peak traffic between groups per unit time

3.2 First step: Minimization of average file download time

At the first step, we focus on the main objective of the P2P file distribution, i.e., minimization of average file download time among peers. The corresponding optimization problem P_1 can be formulated as follows.

$$\min\ \overline{\tau}, \tag{1}$$

$$\text{s.t.}\ z_{0,k,i} = 1, \quad \forall k \in \mathcal{M},\, \forall i \in \mathcal{N}_{\mathrm{D}}, \tag{2}$$

$$z_{0,k,i} = 0, \quad \forall k \in \mathcal{M},\, \forall i \in \mathcal{N}_{\mathrm{P}}, \tag{3}$$

$$x_{t,k,i,j} = \{0,1\}, \quad \forall t \in \mathcal{T}^+,\, \forall i,j \in \mathcal{N},\, i \neq j, \tag{4}$$

$$x_{t,k,i,i} = 0, \quad \forall t \in \mathcal{T}^+,\, \forall k \in \mathcal{M},\, \forall i \in \mathcal{N}, \tag{5}$$

$$x_{t,k,i,j} \leq z_{t-1,k,i},\quad \forall t \in \mathcal{T}^+, \forall k \in \mathcal{M}, \forall i,j \in \mathcal{N}, i \neq j, \tag{6}$$

$$x_{t,k,i,j} \leq 1 - z_{t-1,k,j},\quad \forall t \in \mathcal{T}^+,\, \forall k \in \mathcal{M},\, \forall i,j \in \mathcal{N},\, i \neq j, \tag{7}$$

$$\sum_{j \in \mathcal{N}} x_{t,k,i,j} \leq 1, \quad \forall t \in \mathcal{T}^+, \forall k \in \mathcal{M}, \forall i \in \mathcal{N}, \tag{8}$$

$$x_{t,k,i,j} \leq y_{t-1,i},\ \forall t \in \mathcal{T}^+, \forall k \in \mathcal{M}, \forall i,j \in \mathcal{N}_{\mathrm{P}}, \tag{9}$$

$$\tau_i \leq \tau_j, \quad \forall i,j \in \mathcal{N}_{\mathrm{P}},\, C_i^{\mathrm{out}} \geq C_j^{\mathrm{out}} \tag{10}$$

$$\sum_{k \in \mathcal{M}, j \in \mathcal{N}} (1 - G_{i,j}) x_{t,k,i,j} \leq C_i^{\mathrm{out}}, \forall t \in \mathcal{T}^+, \forall i \in \mathcal{N} \tag{11}$$

$$\sum_{k \in \mathcal{M}, j \in \mathcal{N}} G_{i,j} x_{t,k,i,j} \leq C_i^{\mathrm{in}},\ \forall t \in \mathcal{T}^+, \forall i \in \mathcal{N}_{\mathrm{P}}, \tag{12}$$

$$\sum_{k \in \mathcal{M}} (x_{t,k,i,j} - x_{t,k,j,i}) \leq M G_{i,j},\ \forall t \in \mathcal{T}^+,\, \forall i,j \in \mathcal{N}_{\mathrm{P}}. \tag{13}$$

The objective function (1) indicates the minimization of the average file download time, $\overline{\tau}$, which is given by

$$\overline{\tau} = N_{\mathrm{P}}^{-1} \sum_{i \in \mathcal{N}_{\mathrm{P}}} \tau_i,$$

where τ_i denotes the file download time of peer i and is equivalent to the duration of being a leecher as follows:

$$\tau_i = \sum_{t \in \mathcal{T}} y_{t,i} \quad \forall i \in \mathcal{N}_{\mathrm{P}}.$$

The constraints (2) through (10) are the same as those in case of the strict TFT-based P2P file distribution [13]. The constraints (2) and (3) define the initial piece possession of nodes at $t = 0$: The server has all pieces while all peers have no piece. The constraint (4) gives the range of the decision variable $x_{t,k,i,j}$ and the self loop is prohibited by the constraint (5). The constraint (6) means the feasibility of sending piece k from node i to others while the constraint (7) indicates the need for peer j to retrieve piece k. As in the constraint (8), peer i retrieves piece k from at most one node. The constraint (9) presents the seed departure scenario where peers are selfish and leave the system just after the completion of file downloading. The constraint (10) defines the distribution policy that ensures the peers finish file downloading in descending order of their inter-group upload capacity.

The constraint (11) indicates that each node i can send at most C_i^{out} pieces to peers in different groups at each time. Similarly, the constraint (12) is the constraint for the intra-group communication. The constraint (13) indicates the TFT strategy with the proximity-based relaxation between peer i and peer j. The left hand side of the constraint (13) is the difference of the numbers of pieces exchanged between peer i and peer j at time t. The right hand side of the constraint (13) become zero if peers i and j belong to the different group, i.e., $G_{i,j} = 0$. In this case, peers i and j have to exchange the same number of pieces, that is the TFT constraint is satisfied. On the other hand, if peers i and j belong to the same group, i.e., $G_{i,j} = 1$, the right hand side of the constraint (13) becomes M. Since M is the number of pieces, the constraint (13) is always satisfied, that is the TFT constraint is relaxed according to the proximity between them. Note that the constraint(13) is applied to each pair of peers while the server altruistically provides pieces with peers.

We observe that all the variables are binary or integer and all the equations are linear except $y_{t,i}$ in the constraint (**??**). $y_{t,i}$ includes the nonlinear term of the product of binary variables, i.e., $z_{t,k,i}$, but it can be linearlized [7,13]. As a result, the problem P_1 can be formulated as ILP.

3.3 Second step: Minimization of average inter-group traffic

By solving the first problem P_1, there may be more than one optimal solutions, each of which minimizes the average file download time among peers.

At the second step, we further aim to minimize the average inter-group traffic among peers, which will contribute to reduce the communication costs of peers, e.g., usage-based billing for the Internet connections. The second optimization problem P_2 can be established by modifying P_1 as follows. First, the objective function is replaced with

$$\min \quad \overline{\eta}. \tag{14}$$

The objective function (14) indicates the minimization of the average inter-group traffic among peers, $\overline{\eta}$, which is given by

$$\overline{\eta} = N_{\mathrm{P}}^{-1} \sum_{t \in \mathcal{T}^{+}, k \in \mathcal{M}, i,j \in \mathcal{N}_{\mathrm{P}}} (1 - G_{i,j}) x_{t,k,i,j}. \tag{15}$$

To ensure the minimum average file download time, the following constraint is required:

$$\overline{\tau} = \overline{\tau}^{*}, \tag{16}$$

where $\overline{\tau}^{*}$ is the minimum average file download time, which can be obtained by solving the P_1.

3.4 Third step: Minimization of peak inter-group traffic

When there are multiple solutions obtained by solving problem P_2, we further aim to minimize the peak traffic between groups, which will lead to network-friendly file distribution by reducing the communication cost of groups, e.g., ISPs. The corresponding optimization problem P_3 can be formulated by modifying P_2 as follows. First, the objective function is replaced with

$$\min \quad \eta_{\mathrm{peak}},$$

where η_{peak} is the peak traffic that is the maximum number of pieces exchanged between groups per unit time, and thus η_{peak} should satisfy the following constraint:

$$\sum_{k \in \mathcal{M}, i,j \in \mathcal{N}_{\mathrm{P}}} (1 - G_{i,j}) x_{t,k,i,j} \leq \eta_{\mathrm{peak}}, \quad t \in \mathcal{T}^{+}. \tag{17}$$

The objective function (17) represents the total number of pieces exchanged between groups at time t should be equal or less than than η_{peak}. In addition, we also need to ensure that the second objective should be satisfied:

$$\overline{\eta} = \overline{\eta}^{*},$$

where $\overline{\eta}^{*}$ is the minimized total traffic between groups, which can be obtained by solving the P_2.

Table 2 Default parameter settings.

Parameter	Value
Number of servers, N_{D}	1
Number of peers, N_{P}	8
Number of pieces, M	8
Number of groups, G	2
Upload capacity of server 0, C_0^{out}	2
Upload capacity of peer i to non-group members, C_i^{out}	2
Upload capacity of peer i to group members, C_i^{in}	∞

4 Numerical evaluation

4.1 Evaluation scenario

In this paper, we focus on the analysis of the optimal piece flow under the TFT-based P2P file distribution with the proximity-based TFT relaxation. Since the optimization problems P_1, P_2, and P_3 are ILP as mentioned in Section 3, we can directly solve them using the existing solver CPLEX when the system scale is relatively small. CPLEX solves the ILP by using the combination of branch-and-cut based algorithm and linear programming relaxation, which searches for the optimal solution by updating upper bound (best feasible solution) v_{upper} and lower bound (LP-relaxed best solution) v_{lower}. It can also control the tradeoff between computation time and optimality by introducing allowable tolerance $g_{\max}$ of gap between v_{lower} and v_{upper}. In what follows, $g_{\max}$ is basically set to be zero.

Considering the parameter settings used in the previous work [13, 22, 23], we prepare a default evaluation scenario as follows. Suppose that one server distributes eight pieces to eight peers, i.e., $(N_{\mathrm{D}} = 1, N_{\mathrm{P}} = 8, M = 8)$. In general, the inter-group upload capacity of the server 0, C_0^{out}, tends to be a bottleneck with increase of the number of peers, i.e., $C_0^{\mathrm{out}} \ll N_{\mathrm{P}}$, and thus C_0^{out} is set to be 2. To focus on the fundamental characteristics of the proximity-based TFT relaxation, we consider two groups $(G = 2)$ with the same number of homogeneous peers $i \in \mathcal{N}_{\mathrm{P}}$ whose inter-group upload capacity is identical, i.e., $C_i^{\mathrm{out}} = C_0^{\mathrm{out}} = 2$. The intra-group upload capacity of peer $i \in \mathcal{N}_{\mathrm{P}}$, C_i^{in}, is set to be large enough not to be the bottleneck of the file distribution. The default parameter settings are described in Table 2.

In what follows, we evaluate the relaxed TFT model as well as two existing models (i.e., strict TFT model and no TFT model). Table 3 summarizes the definitions of these three models. Note that the strict TFT model (resp. no TFT model) can also be regarded as a special case of the relaxed TFT model where the number of groups, G, is equal to the number of peers, N_{P}, (resp. 1).

Table 3 Evaluation models.

Communication Model	Intra-group	Inter-group
Relaxed TFT model	w/o TFT	with TFT
Strict TFT model	with TFT	with TFT
No TFT model	w/o TFT	w/o TFT

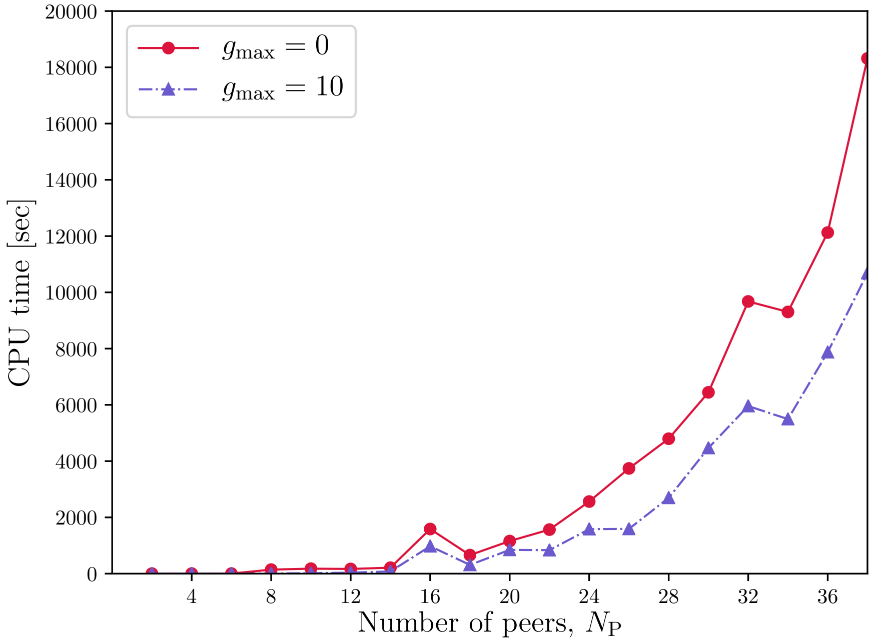

Fig. 1 N_{P} vs CPU time (relaxed TFT model).

4.2 Impact of system scale on CPU time

We first demonstrate the computation complexity of the three-step ILP. Fig. 1 shows the relationship between the number of peers, N_{P}, and CPU time when $g_{\max} = 0, 10$. Since ILP is NP-complete, we observe that the CPU time exponentially increases with the number of peers. We also find that the increase of $g_{\max}$, higher tolerance for the optimality, has a limited impact on reducing the computation time. These results indicate that the ILP approach is difficult to analyze the performance of large-scale systems directly.

To tackle this scalability problem, the clustering approach is applicable as in [23]. In the clustering approach, the whole system is considered as the collection of multiple isolated swarms of peers, i.e., clusters. If the clustered system achieves the optimal average file download time competitive with the non-clustered system under the same system scale, we can estimate the performance of larger-scale systems from the analytical results for small-scale sys-

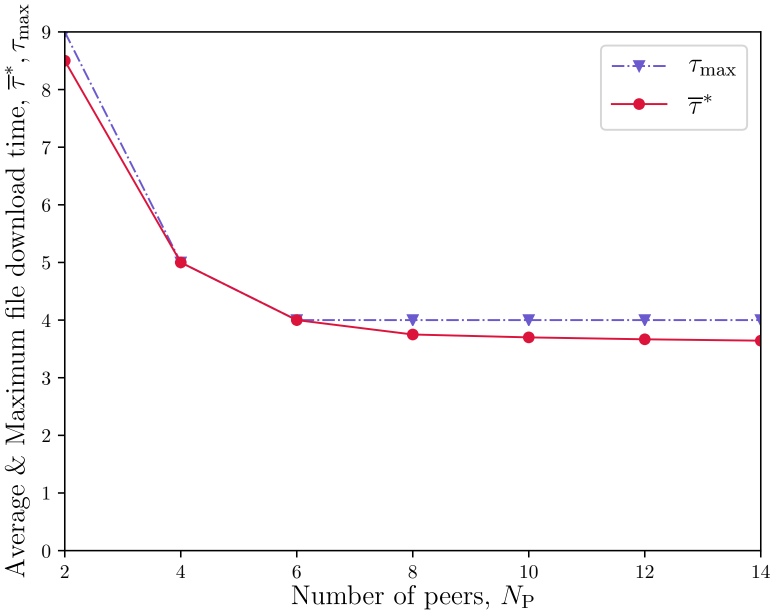

Fig. 2 Impact of system scale on average file download time where the whole system consists of homogeneous group(s) of two peers (relaxed TFT model, $C_0^{\mathrm{out}} = G$).

tems. Fig. 2 presents the relationship between the optimal average/maximum file download time and the system scale where the whole system consists of one or more homogeneous groups of two peers and the server's upload capacity, C_0^{out}, is set to be equal to the number of groups, G. We observe that the increase of the system scale contributes to reducing the optimal average file download time because of the abundant opportunities for inter-cluster communications. At the same time, we can also confirm that the degree of improvement gradually decreases, from which we can expect that larger-scale systems will show the similar performance.

4.3 Impact of system scale on average file download time

In this section, we focus on how the system scale, i.e., the number of peers, N_{P}, and that of pieces, M, affects the average file download time in each model. Fig. 3 shows the transition of the average file download time $\overline{\tau}^*$ with the increase of the number of peers N_{P} when $M = 8$ and $G = 2$. We first observe that the relaxed TFT model shows the same performance as no TFT model, and thus it achieves the lower bound. This desirable phenomena come from the following optimal piece flow. At first, server 0 with $C_0^{\mathrm{out}} = 2$ distributes two different pieces b_1 and b_2 to the two peers k and l, each of which belongs

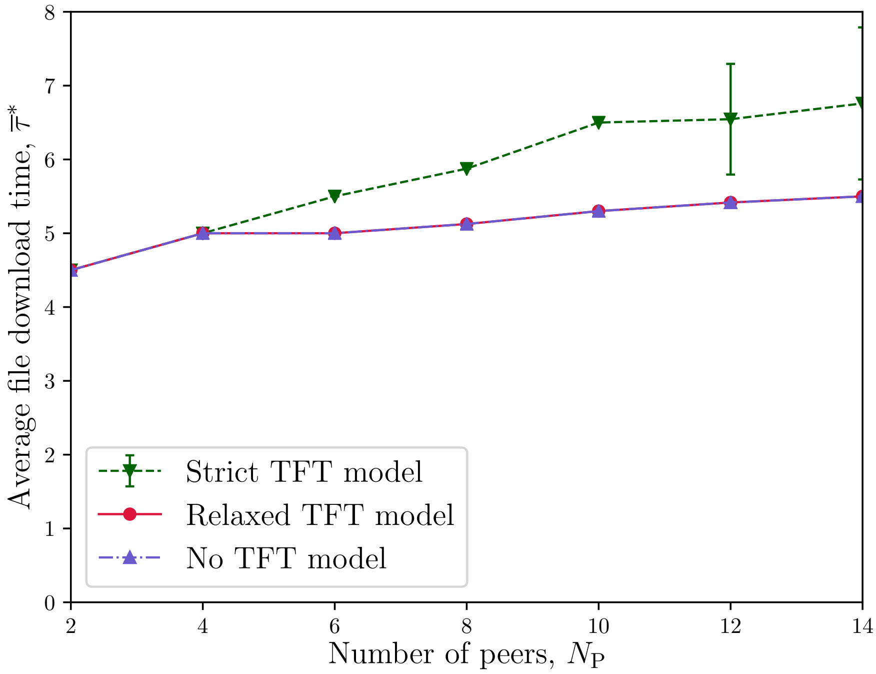

Fig. 3 Impact of the number of peers on average file download time.

to a different group, at $t = 1$. At the next time step $t = 2$, peers k and l speedily transfer b_1 and b_2 to their group members with sufficiently large intra-group upload capacity, i.e., $C_k^{\text{in}} = C_l^{\text{in}} = 10$. At the same time, they can also exchange b_1 and b_2 with each other through inter-group communication with the upload capacity $C_k^{\text{out}} = C_l^{\text{out}} = 2$. Note that server 0 can also send the next two different pieces to the two groups at $t = 2$. As a result, at $t = 3$, two groups can effectively exchange the pieces b_1 and b_2 because of the abundant pairs of peers satisfying the TFT constraint. The optimal piece flow emerges from the repetition of this sustainable cycle.

This can also be confirmed by observing the traffic pattern in case of the relaxed TFT model. Fig. 4 shows how the traffic pattern changes with increase of the number of peers. We first observe that both intra-group and inter-group traffic grows with increase of the number of peers while the server traffic does not almost change. In other words, the relaxed TFT model can achieve the effective file distribution where the upload capacity of the server is fully utilized and piece distribution through peers are also encouraged. In particular, we can confirm that the intra-group traffic much contributes to the file distribution compared with the inter-group traffic and server traffic.

Next, we focus on the fact that the average file download time of the strict TFT model, which is the upper bound, faster grows with the increase of the number of peers, compared with other two models. This is because peer

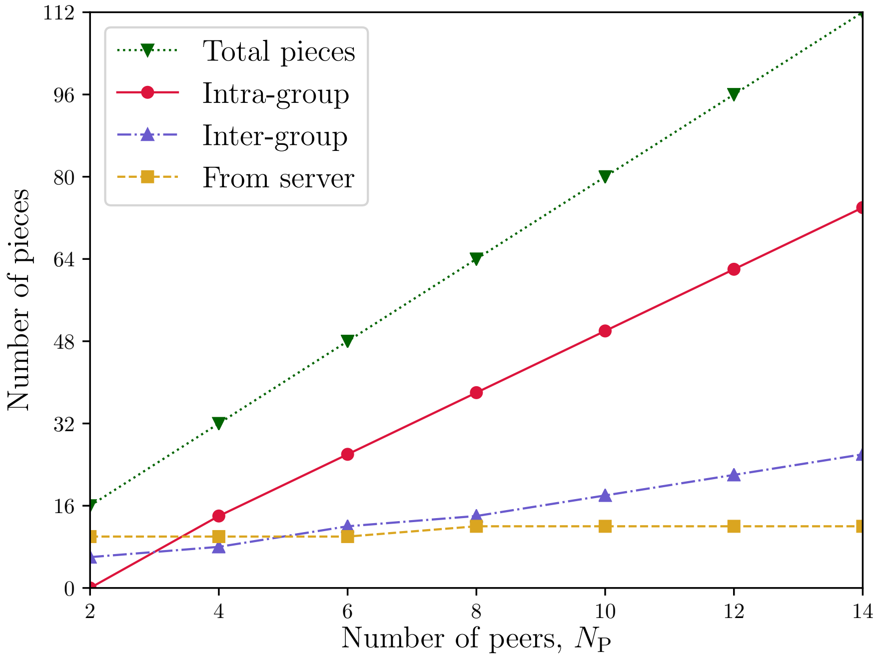

Fig. 4 Relationship between the number of peers and traffic pattern (relaxed TFT model).

cannot have enough pieces to exchange with others, due to the bottleneck of the inter-group upload capacity of server 0.

Finally, we focus on the impact of the number of pieces M on the average file download time. Fig. 5 illustrates the relationship between the number of pieces, M, and the average file download time $\overline{\tau}^*$, in case of the relaxed TFT model with $N_\mathrm{P} = 8$ and $G = 2$. We observe the average file download time linearly increases with M, i.e., the content size. This indicates that the above-mentioned sustainable cycle of piece exchange can be achieved in the relaxed TFT model.

4.4 Impact of number of groups on average file download time

As mentioned in Section 4.1, the strict TFT model and no TFT model are special cases of the relaxed TFT model, depending on the number of groups, G. In other words, the system performance of the relaxed TFT model will change according to G. Fig. 6 depicts the relationship between the number of groups, G, and the average file download time $\overline{\tau}^*$ when $N_\mathrm{P} = 8, M = 8$, and the inter-group upload capacity of the server, C_0^{out}, increases. We set $G = \{1, 2, 4, 8\}$ where each group size is kept to be identical. For instance, the number of peers in each group becomes two when $G = 4$.

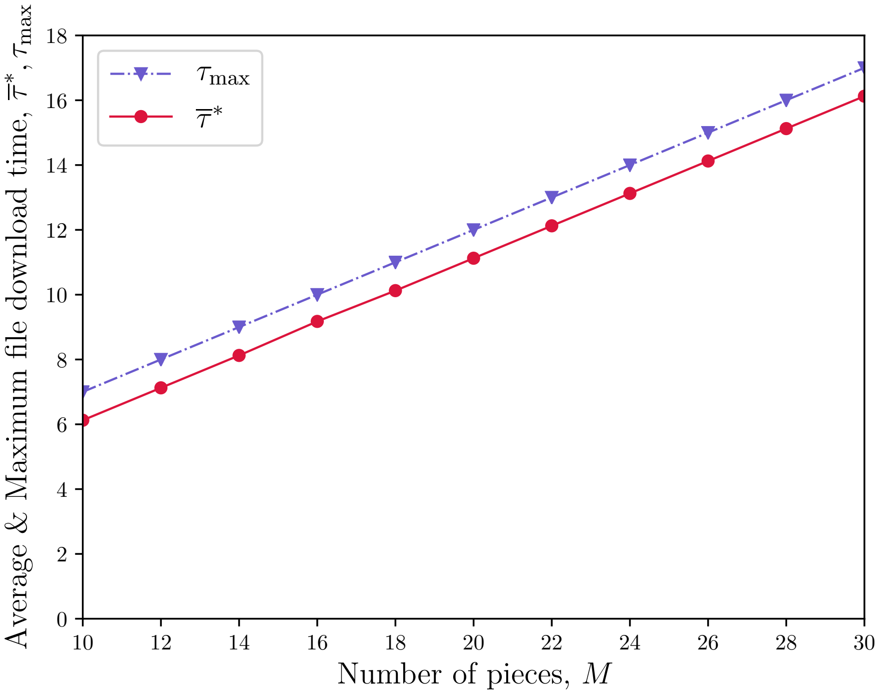

Fig. 5 Impact of the number of pieces on average file download time (relaxed TFT model).

As mentioned in Section 4.1, the relaxed TFT model with $G = N_{\mathrm{P}} = 8$ (resp. $G = 1$) is equivalent to the strict TFT model (resp. no TFT model). As we expect, the average file download time of the relaxed TFT model grows with the increase of the number of groups G. However, we also observe that the results of $G = 2$ and $G = 4$ can be competitive with those of $G = 1$, that is the lower bound, if the inter-group upload capacity of the server 0, C_0^{out}, is relatively small, e.g., $C_0^{\mathrm{out}} = [1, 3]$. In actual systems, the inter-group upload capacity of the server tends to be the bottleneck, and thus we can expect that the TFT constraint in the inter-group communication does not so much degrade the average file download time, with the help of the speedy intra-group communication. The TFT constraint in the inter-group communication can contribute to establishing the relationship of mutual trust between anonymous peers in different groups.

4.5 Impact of upload capacity of server

Increasing the server's upload capacity will contribute to speedy file distribution but it also requires the investment cost. Fig. 7 illustrates the relationship between the server's upload capacity, C_0^{out} and average file download time when the file size, i.e., the number of pieces, M, is set to be 10, 20, and 30,

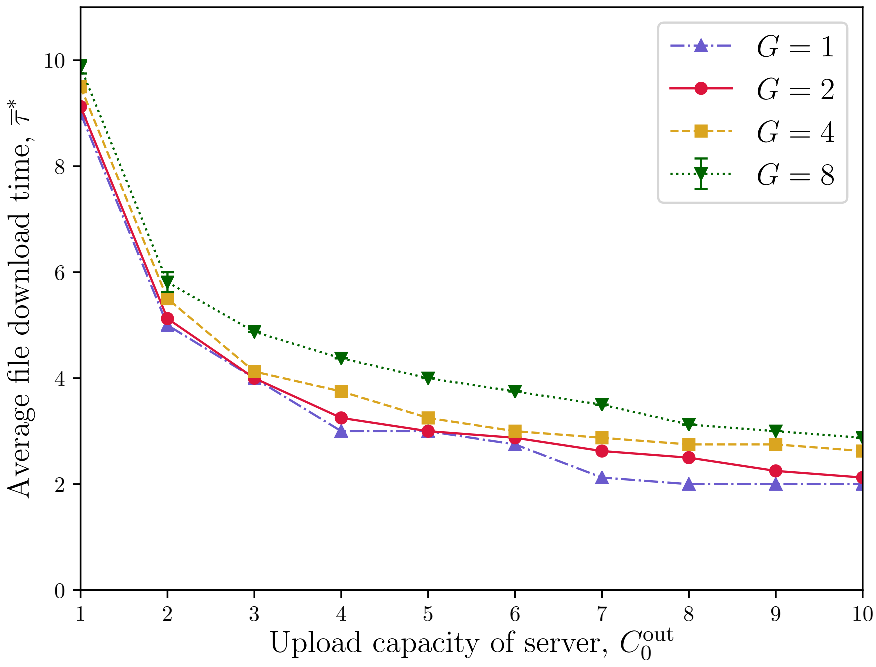

Fig. 6 Impact of the number of groups on average file download time (relaxed TFT model).

respectively. Note that the average file download time is normalized by that in case of $C^{\text{out}} = 1$ to compare the results among different scenarios. In addition, both the intra-group and inter-group upload capacity of peers are set to be sufficiently large values, i.e., $C_i^{\text{out}} = C_i^{\text{in}} = \infty$, such that the server's upload capacity becomes the bottleneck. The impact of the upload capacity of peers will be discussed in Section 4.6. As we expect, the performance improves with increase of C_0^{out}. In particular, relatively small C_0^{out} can achieve much performance improvement, e.g., 71.4–76.7% reduction of the average file download time in case of $C_0^{\text{out}} = 5$. The larger the file size, M, is, the more the reduction degree of average file download time becomes.

4.6 Impact of upload capacity of peers

In Section 4.5, we demonstrated how the optimal average file download time changes with the server's upload capacity C_0^{out} when both the intra-group and inter-group upload capacities of peers are abundant. Since the inter-group upload capacity tends to be much smaller than the intra-group upload capacity, we next focus on the case where the inter-group upload capacity becomes the bottleneck.

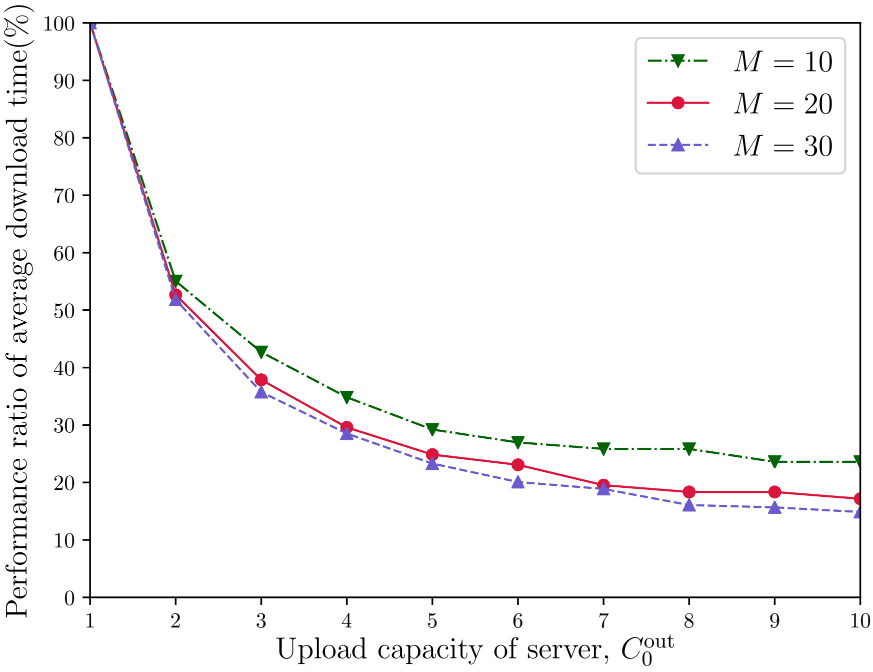

Fig. 7 Impact of upload capacity of server on average file download time (relaxed TFT model).

Fig. 8 illustrates the relationship between the average file download time and server's upload capacity C_0^{out} when the inter-group upload capacity of peers, i.e., C_i^{out}, is set to be $1, 2, 4$, and ∞. Note that the scenario with $C_i^{\mathrm{out}} = \infty$ is the same one used in Section 4.5. We observe that small inter-group upload capacity of peers, e.g., $C_i^{\mathrm{out}} = 2$, can achieve almost the optimal system performance especially in case of relatively small server's upload capacity, e.g., $C_0^{\mathrm{out}} = [1, 6]$.

4.7 Impact of system scale on average inter-group traffic

Next, we focus on the impact of the system scale on the average inter-group traffic among peers, $\overline{\eta}$, which corresponds to the second objective. Fig. 9 depicts the transition of the average inter-group traffic among peers with increase of the number of the peers, N_{P}, when $M = 8$ and $G = 2$. Note that we only compare the results of the relaxed TFT model with those of no TFT model by considering the fact that they achieve the same performance in terms of the average file download time, as shown in Section 4.3. We observe that the gap of $\overline{\eta}$ between relaxed TFT model and no TFT model expands with the system scale. Since the server's upload capacity C_0^{out} is fixed to be 2, the increase

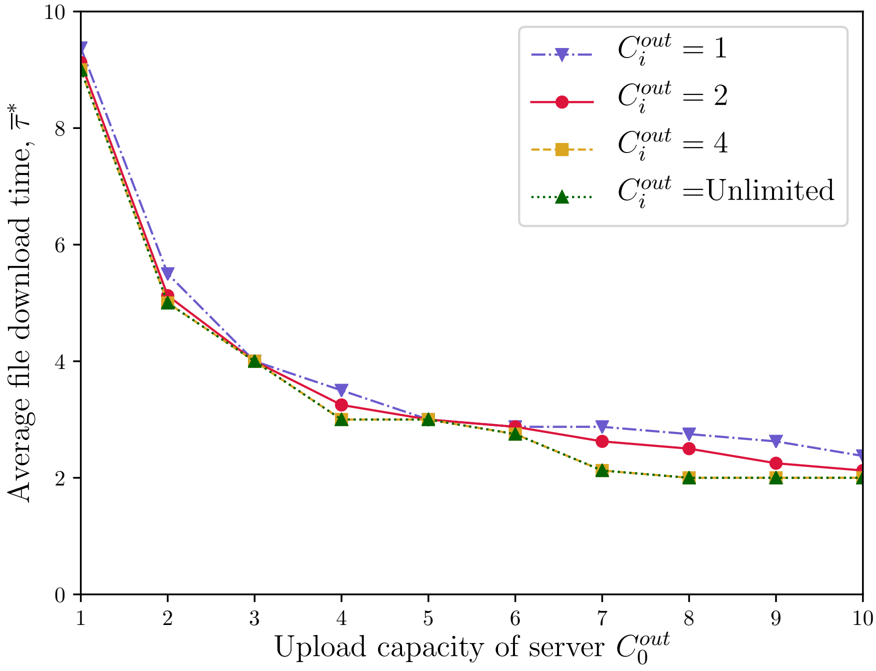

Fig. 8 Impact of upload capacity of peers on average file download time (relaxed TFT model).

of peers requires more piece distribution among peers to reduce the average file download time. In addition, the relaxed TFT model still needs the TFT constraint on the inter-group communication. As a result, the relaxed TFT model yields more inter-group traffic than no TFT model.

4.8 Impact of system scale on peak inter-group traffic

Finally, we analyze the impact of system scale on the peak traffic between groups, η_{peak}, which corresponds to the third objective. Fig. 10 shows the relationship between the number of peers, N_{P}, and the peak traffic between groups, η_{peak}, in case of $M = 8$ and $G = 2$. In Fig. 10, we also only compare the relaxed TFT model with no TFT model. We observe that the results of the relaxed TFT model approach to be twice as those of no TFT model when the number of peers increases, i.e., $N_{\text{P}} \geq 6$. This is because the relaxed TFT model requires the equivalent piece exchange in the inter-group communication, which is not necessarily required in no TFT model.

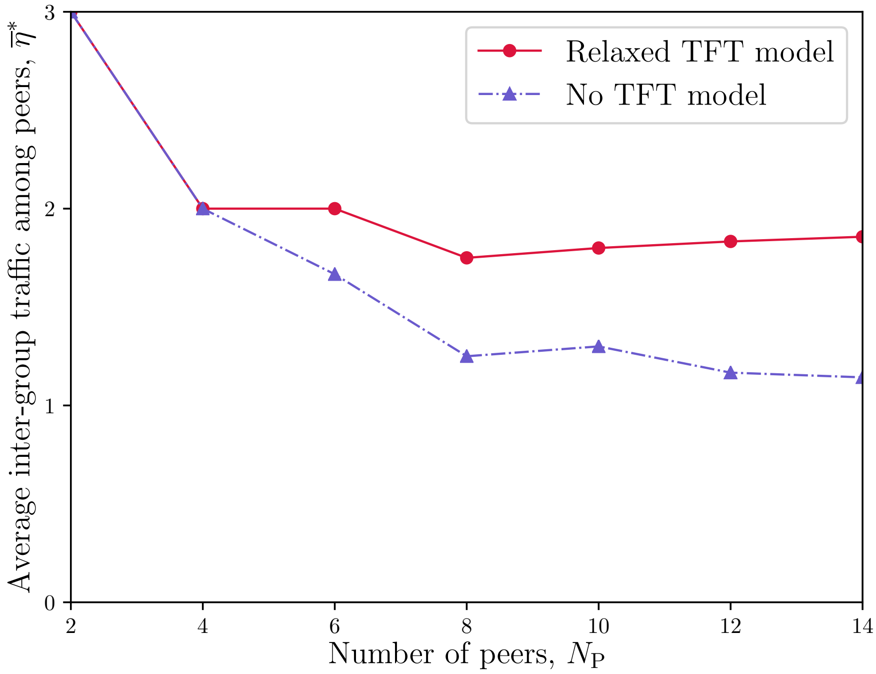

Fig. 9 Impact of the number of peers on average inter-group traffic among peers.

5 Conclusion

In this paper, we have first modeled the locality-aware TFT-based P2P file distribution where only inter-group communications are governed by the TFT constraint. We have formulated the optimal piece flow determination problem as the third-step ILP, which minimizes the average file download time among peers, the average inter-group traffic among peers, and the peak traffic among groups, in this order. Through the numerical results, we mainly have shown the following characteristics. (1) The relaxed TFT model shows the competitive performance with no TFT model in terms of the average file download time when the upload capacity of the server is a bottleneck, with the help of altruistic intra-group communications. (2) The upload capacity of the server has much larger impact on the average file download time than that of peers but a relatively small upload capacity of the server can drastically shorten the average file download time. (3) The relaxed TFT model with a moderate number of groups can achieve better average file download time than the strict TFT model. (4) Both the average and peak inter-group traffic tend to be about twice as those of no TFT model, due to its TFT constraint on inter-group communication.

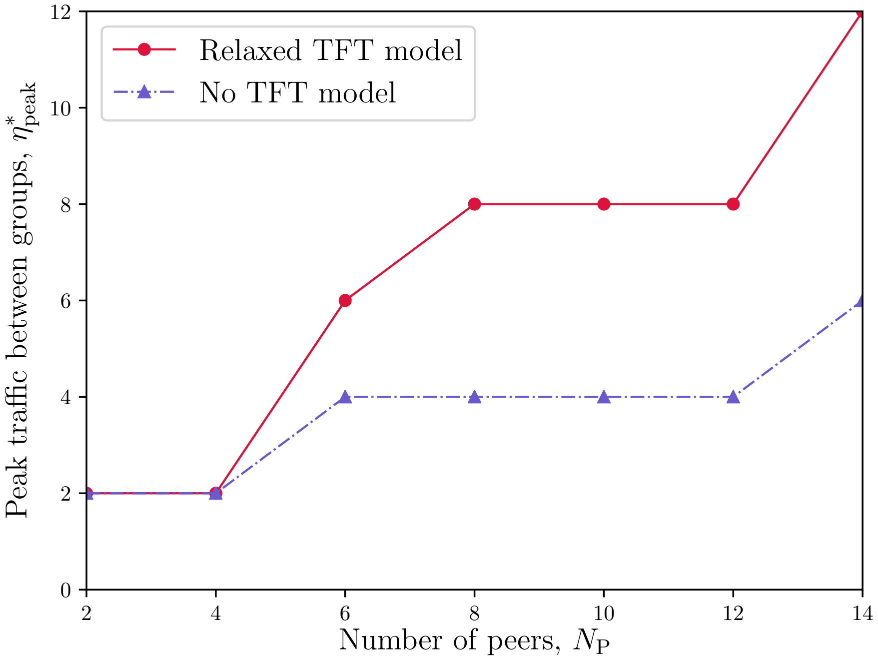

Fig. 10 Impact of the number of peers on peak traffic between groups.

Acknowledgements This work was supported in part by KDDI Foundation Research Grant Program, JSPS KAKENHI (A) under Grant 19H01103, and JSPS KAKENHI (C) under Grant 19K11942.

References

1. BitTorrent. available at `https://www.bittorrent.com`
2. Deploy and Update Windows10. available at `https://docs.microsoft.com/en-us/windows/deployment/`
3. IBM ILOG CPLEX Optimizer. available at `https://www.ibm.com/analytics/cplex-optimizer`
4. Azzedin, F., Yahaya, M.: Modeling BitTorrent Choking Algorithm Using Game Theory. Future Generation Computer Systems **55**(2), 255–265 (2016)
5. Blond, S.L., Legout, A., Dabbous, W.: Pushing BitTorrent Locality to the Limit. Computer Networks **55**(3), 541–557 (2011)
6. Carbunaru, C., Teo, Y.M., Leong, B., Ho, T.: Modeling Flash Crowd Performance in Peer-to-Peer File Distribution **25**(10), 2617–2626 (2014)
7. Chen, D., Batson, R.G., Dang, Y.: Applied Integer Programming. Wiley (2010)
8. Cuevas, R., Laoutaris, N., Yang, X., Siganos, G., Rodriguez, P.: BitTorrent Locality and Transit TrafficReduction: When, Why, and at What Cost? IEEE Transactions on Parallel and Distributed Systems **25**(5), 1177–1189 (2014). DOI 10.1109/TPDS.2013.109
9. D'Alessandro Costa, M.A., Gonçalves Rubinstein, M.: Performance Analysis of a Locality-Aware BitTorrent Protocol in Enterprise Networks. Peer-to-Peer Networking and Applications pp. 1–12 (2018)

10. Ferragut, A., Paganini, F.: Fluid Models of Population and Download Progress in P2P Networks. IEEE Transactions on Control of Network Systems **3**(1), 34–45 (2016). DOI 10.1109/TCNS.2015.2434092
11. Ghaderzadeh, A., Kargahi, M., Reshadi, M.: InFreD: Intelligent Free Rider Detection in Collaborative Distributed Systems. Journal of Network and Computer Applications **78**, 134 – 145 (2017). DOI https://doi.org/10.1016/j.jnca.2016.11.007
12. Goldenberg, D.K., Qiuy, L., Xie, H., Yang, Y.R., Zhang, Y.: Optimizing Cost and Performance for Multihoming. In: Proc. of the ACM SIGCOMM Computer Communication Review, pp. 79–92 (2004)
13. Hasegawa, M., Sasabe, M., Takine, T.: Analysis of Optimal Scheduling in Tit-for-Tat-Based P2P File Distribution. IEICE Transactions on Communications **97**(12), 2650–2657 (2014)
14. Huang, W., Wu, C., Li, Z., Lau, F.C.: The Performance and Locality Tradeoff in Bittorrent-like File Sharing Systems. Peer-to-Peer Networking and Applications **7**(4), 469–484 (2014). DOI 10.1007/s12083-012-0190-2
15. Jin, X., Kwok, Y.K.: Coercion Builds Cooperation in Dynamic and Heterogeneous P2P Live Streaming Networks. Computer Networks **81**, 1 – 18 (2015). DOI https://doi.org/10.1016/j.comnet.2015.02.006
16. Jun, S., Ahamad, M.: Incentives in BitTorrent Induce Free Riding. In: Proceedings of the 2005 ACM SIGCOMM Workshop on Economics of Peer-to-peer Systems, P2PECON '05, pp. 116–121. ACM, New York, NY, USA (2005). DOI 10.1145/1080192.1080199
17. Kang, X., Wu, Y.: Incentive Mechanism Design for Heterogeneous Peer-to-Peer Networks: A Stackelberg Game Approach. IEEE Transactions on Mobile Computing **14**(5), 1018–1030 (2015). DOI 10.1109/TMC.2014.2343628
18. Lu, K., Wang, S., Xie, L., Wang, Z., Li, M.: A Dynamic Reward-Based Incentive Mechanism: Reducing the Cost of P2P systems. Knowledge-Based Systems **112**, 105 – 113 (2016). DOI https://doi.org/10.1016/j.knosys.2016.09.002
19. Magharei, N., Rejaie, R., Rimac, I., Hilt, V., Hofmann, M.: ISP-Friendly Live P2P Streaming **22**(3), 244–256 (2014)
20. Nishi, Y., Sasabe, M., Kasahara, S.: Impact of Locality-awareness on Tit-for-Tat-based P2P File Distribution. In: to be presented at IEEE Consumer Communications & Networking Conference, pp. 1–6 (2020)
21. Pouwelse Johanand Garbacki, P., Epema, D., Sips, H.: The Bittorrent P2P File-Sharing System: Measurements and Analysis. In: M. Castro, R. van Renesse (eds.) Peer-to-Peer Systems IV, pp. 205–216. Springer Berlin Heidelberg, Berlin, Heidelberg (2005)
22. Sasabe, M.: Analysis of Optimal Piece Flow in Tit-for-Tat-Based P2P Streaming. Computer Networks **139**(7), 60–69 (2018)
23. Sasabe, M.: Topological Influence on Optimality of Tit-for-Tat based P2P Content Distribution. Peer-to-Peer Networking and Applications pp. 1–12 (2019)
24. Seibert, J., Torres, R., Mellia, M., Munafo, M.M., Nita-Rotaru, C., Rao, S.: The Internet-Wide Impact of P2P Traffic Localization on ISP Profitability. IEEE/ACM Transactions on Networking **20**(6), 1910–1923 (2012). DOI 10.1109/TNET.2012.2190093
25. de Souza e Silva, E., Leão, R.M., Menasché, D.S., Towsley, D.: On the Scalability of P2P Swarming Systems. Computer Networks **151**(3), 93–113 (2019)
26. Wang, H., Liu, J.: Exploring Peer-to-Peer Locality in Multiple Torrent Environment. IEEE Transactions on Parallel and Distributed Systems **23**(7), 1216–1226 (2012). DOI 10.1109/TPDS.2011.253
27. Wu, T.Y., Lee, W.T., Guizani, N., Wang, T.M.: Incentive Mechanism for P2P File Sharing Based on Social Network and Game Theory. Journal of Network and Computer Applications **41**, 47 – 55 (2014). DOI https://doi.org/10.1016/j.jnca.2013.10.006
28. Xia, R.L., Muppala, J.K.: A Survey of BitTorrent Performance **12**(2), 140–158 (2010)
29. Zhang, X., Wang, N., Cao, Y., Peng, L., Meng, H.: A Stochastic Analytical Modeling Framework on ISP–P2P Collaborations in Multidomain Environments **12**(3), 2320–2331 (2018)
30. Zhao, J., Wu, C., Lin, X.: Locality-aware Streaming in Hybrid P2P-Cloud CDN Systems. Peer-to-Peer Networking and Applications **8**(2), 320–335 (2015). DOI 10.1007/s12083-013-0233-3