Submitted version

Noname manuscript No.
(will be inserted by the editor)

Topological Influence on Optimality of Tit-for-Tat Based P2P Content Distribution

Masahiro Sasabe

Received: date / Accepted: date

Abstract Peer-to-Peer (P2P) content distribution is a powerful scheme to distribute content on the Internet. Since the P2P content distribution relies on the cooperation among peers, one of the most famous P2P file distribution systems, BitTorrent, has applied a game theoretical approach called the Tit-for-Tat (TFT) strategy to encourage selfish peers to cooperatively exchange the fragments of file, i.e., pieces, with others. In recent years, the basic characteristics of such TFT-based P2P content distribution under a full-mesh network has been investigated through modeling the determination of optimal piece flow as Integer Linear Programming (ILP). However, the topological influence on optimal flow has not been revealed yet. In this paper, we propose an approach to analyze the topological influence by extending the previous model. Through numerical results, we reveal that the optimal piece flow can be achieved on a hierarchical and circular topology with $O(N_\mathrm{P})$ links where N_P is the number of peers. We also show the whole network can be divided into multiple sub-networks while keeping the system performance.

Keywords Peer-to-Peer (P2P) · Content distribution · Tit-for-Tat strategy · Integer Linear Programming (ILP) · Optimal piece flow · Topological influence

---

This research was partly supported by KDDI Foundation Research Grant Program, Japan.

Masahiro Sasabe
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: sasabe@is.naist.jp

1 Introduction

With the proliferation of broadband Internet access, content distribution has been growing for various types of content: disk images of operating systems (Linux, FreeBSD, etc.), software and its security patches, and video/audio streaming media. Most of them still rely on the conventional client-server architecture, where a server or a limited number of servers serve such content to many clients. Although the client-server architecture is suitable for user management, it potentially has a drawback of scalability to the number of clients, due to the bottleneck of the upload capacity of servers. When new versions of software and security patches or latest videos are released, many users will simultaneously access to the servers and cuases Flash Crowd [8]. Peer-to-Peer (P2P) content distribution systems have been expected to solve such scalability problems. For example, BitTorrent [2], which is one of the most famous P2P file distribution systems, has been used for distributing the disk images of operating systems, and Windows 10 applies P2P file distribution called Windows Update Delivery Optimization [11] to distribute security patches. PPTV [14] is one the most successful P2P streaming services.

In P2P content distribution systems, the server divides content into small fragments called pieces and distributes them to clients. The clients that retrieved those pieces can also transfer them to other clients. This indicates that the clients can also play the role of servers, and thus they are called peers. With the help of honest contribution of peers, the P2P content distribution systems will overcome the drawback of the conventional client-server content distribution systems in terms of single point of failure and lack of scalability.

In the P2P content distribution, however, peers are required to contribute high-rate and/or long-term data transfer to others, which may decrease their quality of other network services being enjoyed. As a result, it has been pointed out that many peers behave as free riders, which hesitate about transferring data to others, in actual systems [7]. To tackle this problem, BitTorrent applies a game theoretical approach called Tit-for-Tat (TFT) strategy. In the original TFT strategy, a player initially takes a cooperative action and then imitates the action taken by the opponent. Therefore, the cooperative relationship between them is kept unless the opponent defects. In case of P2P file distribution systems, cooperation (resp. noncooperation) corresponds to acceptance (resp. denial) of requests from others. As a result, each peer has to send pieces to others so as to retrieve its demanding pieces from them.

Recently, the optimality of such TFT-based P2P content distribution, i.e., file distribution and streaming, has been analyzed [5,16]. The authors develop a discrete-time model of TFT-based P2P file distribution. They formulate finding its optimal piece flow as Integer Linear Programming (ILP), where the objective function is minimizing the average download time (resp. play-out delay) among peers in case of file distribution (resp. streaming distribution). Note that the piece flow describes how and when each piece is transferred among nodes, i.e., servers or peers. They can obtain the optimal piece flow using the existing solver, i.e., CPLEX [1]. By analyzing the optimal piece flow, they reveal piece sending/receiving strategies to yield the optimal piece flow.

In [5,16], the authors implicitly assume that the P2P network is a full-mesh network, where each peer can freely exchange pieces with others under the TFT constraint. In actual systems, it is desirable that peers are sparsely connected with others to reduce maintenance/communication overhead and improve dispersibility of the system. In this paper, we first try to reveal an optimal topology for TFT-based P2P content distribution, which can achieve the optimal performance with the minimum number of links. We will see this topology finding problem can also be reduced to ILP as in previous work [5,16]. Through numerical results, we analyze the topological influence on TFT-based P2P file distribution and tackle the following fundamental questions. *How many links will be required to achieve the optimal distribution? Can the whole network be divided into multiple isolated sub-networks without deteriorating the performance? What structure does the optimal topology have and does the structure change depending on the content type, i.e., file or streaming?*

The main contributions of this paper are as follows:

1. We formulate ILP to find the optimal topology that can achieve the optimal performance, i.e., minimum average file downloading time in file distribution case and minimum average play-out delay in streaming distribution case, with the minimum number of links. By solving the ILP using existing solver, we can obtain numerical examples of optimal topologies and optimal piece flow over them.
2. Through numerical results, we reveal that (1) the topological structure becomes more important in case of file distribution, (2) both file and streaming distribution can achieve the optimal piece flow under the hierarchical and circular topology with a smaller number of links, i.e., $O(N_{\mathrm{P}})$, compared with that of full-mesh network, i.e., $O(N_{\mathrm{P}}^2)$, (3) the whole network can be divided into multiple sub-networks, i.e., clusters, while keeping the system performance, (4) the optimal piece flow over the optimal topology demonstrates that the server's optimal piece sending strategy is the same as that in a full-mesh network and peers can effectively exchange the pieces retrieved from the server with neighboring peers.

The rest of the paper is organized as follows. In Section 2, we review the related work. We present modeling TFT-based content distribution and formulation of its optimal piece flow and topology in Section 3. After analyzing the topological influences on the optimality of TFT-based P2P content distribution in Section 4, Section 5 gives the conclusion.

2 Related Work

There are many studies on topological structures in P2P content distribution systems: file distribution cases [6,22,12,17,23,21,18] and streaming distribution cases [20,19,9]. Most of them are conducted by measurement-based approaches [17,20,18,19,9] while analytical and/or simulation-based approaches [6,22,12,23,21] also exist.

In [17], the authors investigate the characteristics and dynamics of P2P topologies, by crawling the actual Gnutella network. They reveal the small network structure, high resiliency to random peer departure and systematic attacks, and stable core structure consisting of ling-lived peers. In [18], the authors measure the actual BitTorrent topology and reveal that the topology has short distances and low clustering coefficients, and the degree-frequency exhibits a Gaussian-like distribution.

In [20], the authors investigate the topological properties and dynamics of an actual live P2P streaming system, through a long-time observation. They observe that live P2P streaming topology has high scalability, a high level of link reciprocity [4], a clustering structure, and no power-law degree distribution. In [19], the authors measure the topology of a P2P live streaming system, and derive mathematical models for the distribution of node degree. They also reveal the topology is similar to a random graph. In [9], the authors investigate the topology of a P2P VoD system using crawling-based active measurements. They reveal that the topology has the exponential piecewise node degree distribution and small-world property.

In [23], the authors propose a directed and weighted graph model to describe BitTorrent protocol based P2P networks. With the help of complex networks theory based analysis, they find that the node degree, i.e. the number of neighbors, follows a power-law distribution and there is a positive correlation between flow betweenness centrality [15] and node degree. In [21], the authors propose an evolution model of unstructured P2P file-sharing networks, which consists of four evolution events: node addition, node departure, connection establishment, and edge deletion. They also derive a difference equation of degree distribution, which can present both the internal dynamics and external dynamics of network topology. In [10], the authors propose position-based topology that matches P2P network with physical topology, which achieves better characteristics than other complex networks' models like small-world and scale-free networks. In [6], the authors investigates the tradeoff between the performance and locality in BitTorrent-like P2P file sharing systems. In [22,12], the authors model the TFT-based piece exchange as network formation game and clarifies the existence of Nash equilibrium and mechanism for effective bandwidth exchange through theoretical analysis.

There is no study that reveals the topological influences on the optimality of TFT-based P2P file/streaming distribution. In this paper, we tackle this problem by formulating the determination of optimal piece flow and optimal topology in TFT-based P2P content distribution as ILP.

3 Modeling TFT-based P2P content distribution and formulation of its optimal piece flow

In this section, we first develop a discrete-time model of TFT-based P2P content distribution. Next, we formulate its optimal piece flow as ILP by extending our previous work [5,16]. Note that the content can be either files or streaming media. In case of streaming distribution, there are two kinds of streaming media, i.e., stored media and live media. For simplicity, we mainly focus on the stored-media case in the following but we can easily extend the following model to the live-media case as in [16].

3.1 Model

In this paper, TFT-based P2P content distribution is modeled as a discrete-time system. We consider set $\mathcal{N}_{\mathrm{D}} = \{0\}$ of server and that $\mathcal{N}_{\mathrm{P}} = \{1, 2, \dots, N_{\mathrm{P}}\}$ of peers. Note that the server can also be regarded as a virtual server composed of multiple servers. For simplicity of notation, the server and peers are generically called nodes and we define $\mathcal{N} = \mathcal{N}_{\mathrm{D}} \cup \mathcal{N}_{\mathrm{P}}$ and $N = 1 + N_{\mathrm{P}}$. Peers that have completed content retrieving are called seeds. Other peers are called leechers. Suppose that all peers try to download (resp. play-out) a specific content in file (resp. streaming) distribution, where the content consists of set $\mathcal{M} = \{1, 2, \dots, M\}$ of pieces.

Let C_i $(i \in \mathcal{N})$ be the upload capacity of node i. As in [5,16], we assume that C_i is a natural number and the maximum number of pieces that node i can transfer at each time step. Without the loss of generality, we assume that $C_i \geq C_j$ for $i < j \in \mathcal{N}_{\mathrm{P}}$. On the other hand, we assume that there is no limitation of download capacity for each peer by taking account of the asymmetry of uplink and downlink channel speeds of the current Internet connections, e.g., ADSL and cable Internet. Note that the TFT strategy will bound peer i's download speed by the total upload capacity of server and seeds. We assume that piece transfers between nodes are synchronized and completed in a unit time. As a result, each node i can transfer at most C_i pieces at each time step. Note that each node can transfer multiple, different pieces to a specific leecher under the upload capacity constraint.

We define decision variables $x_{t,k,i,j}$ $(t = 1, 2, \dots, T,\ k \in \mathcal{M},\ i, j \in \mathcal{N})$ as

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

where T denotes the maximum time step that ensures the content retrieval of all peers, which can be bounded by the time step required for direct distribution from the server to all peers [5]. For simplicity of description, let $\mathcal{T}$ and $\mathcal{T}^{+}$ be

$$\mathcal{T} = \{0, 1, \dots, T\}, \quad \mathcal{T}^{+} = \{1, 2, \dots, T\},$$

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, $\{1,\dots,N_{\mathrm{P}}\}$
$\mathcal{N}$	The set of nodes, $\{0,1,\dots,N\}$, $\mathcal{N}=\mathcal{N}_{\mathrm{D}}\cup\mathcal{N}_{\mathrm{P}}$
$\mathcal{M}$	The set of pieces, $\{1,2,\dots,M\}$
C_i	Upload capacity of node i
$x_{t,k,i,j}$	Decision variables of piece transfers
$y_{t,i}$	Variables of nodes' roles (1: leechers, 0: seeds or server)
$z_{t,k,i}$	Variables of piece possession (1: possession, 0: missing)
$a_{i,j}$	Decision variables of connection (1: connected, 0: unconnected)
τ_i	Download time of peer i
w_i	Play-out delay of peer i

respectively.

We can keep track of the process of the content distribution through $x_{t,k,j,i}$ by additionally defining $z_{t,k,i}$ ($t \in \mathcal{T}$, $k \in \mathcal{M}$, $i \in \mathcal{N}$) and $y_{t,i}$ ($t \in \mathcal{T}$, $i \in \mathcal{N}$) as

$$z_{t,k,i} = \begin{cases} z_{0,k,i} + \sum_{s=1}^{t} \sum_{j\in\mathcal{N}} x_{s,k,j,i}, & \text{if } i \in \mathcal{N}_{\mathrm{P}}, \\ 1, & \text{otherwise}, \end{cases} \tag{1}$$

$$y_{t,i} = \begin{cases} 1 - \prod_{k\in\mathcal{M}} z_{t,k,i} & \text{if } i \in \mathcal{N}_{\mathrm{P}}, \\ 0, & \text{otherwise}, \end{cases} \tag{2}$$

respectively, where empty sum is defined to be zero. Note that $z_{t,k,i}$ represents the state of piece possession of node i at time step t, i.e., $z_{t,k,i} = 1$ if peer i has piece k at time step t, and otherwise $z_{t,k,i} = 0$. On the other hand, $y_{t,i}$ represents the role of peer i at time step t, i.e., $y_{t,i} = 1$ if peer i is a leecher at time step t, and otherwise $y_{t,i} = 0$, which indicates that peer i is a seed. Note that server ($i \in \mathcal{N}_{\mathrm{D}}$) constantly have all pieces ($z_{t,k,i} = 1, y_{t,i} = 0$). Table 1 summarizes notations we introduced.

3.2 Optimal piece flow without topological constraints

In this section, we first introduce the determination problem of optimal piece flow without topological constraints, which is the exactly same as that in our previous work [5,16]. In what follows, we give the formulation for file distribution in Section 3.2.1 and that for streaming distribution in Section 3.2.2.

3.2.1 File distribution case

From a viewpoint of the entire system (i.e., social optimum), an optimal piece flow in P2P file distribution can be regarded as a piece flow that minimizes the average download time among peers. We thus formulate a minimization problem $\mathrm{P_F}$ of the average download time as follows.

$$\begin{aligned}
\min \quad & \overline{\tau}, && (3)\\
\text{s.t.} \quad & z_{0,k,i} = 1, \quad \forall k \in \mathcal{M}, \forall i \in \mathcal{N}_{\mathrm{D}}, && (4)\\
& z_{0,k,i} = 0, \quad \forall k \in \mathcal{M}, \forall i \in \mathcal{N}_{\mathrm{P}}, && (5)\\
& x_{t,k,i,i} = 0, \quad \forall t \in \mathcal{T}^{+}, \forall k \in \mathcal{M}, \forall i \in \mathcal{N}, && (6)\\
& x_{t,k,i,j} \in \{0,1\}, \quad \forall t \in \mathcal{T}^{+}, \forall k \in \mathcal{M}, \forall i,j \in \mathcal{N}, i \neq j, && (7)\\
& x_{t,k,i,j} \leq z_{t-1,k,i}, \ \forall t \in \mathcal{T}^{+}, \forall k \in \mathcal{M}, \forall i,j \in \mathcal{N}, \ i \neq j, && (8)\\
& x_{t,k,i,j} \leq 1 - z_{t-1,k,j}, \quad \forall t \in \mathcal{T}^{+}, \forall k \in \mathcal{M}, \forall i,j \in \mathcal{N}, \ i \neq j, && (9)\\
& \sum_{j \in \mathcal{N}} \sum_{k \in \mathcal{M}} x_{t,k,i,j} \leq C_i, \quad \forall t \in \mathcal{T}^{+}, \forall i \in \mathcal{N}, && (10)\\
& \sum_{j \in \mathcal{N}} x_{t,k,j,i} \leq 1, \quad \forall t \in \mathcal{T}^{+}, \forall k \in \mathcal{M}, \forall i \in \mathcal{N}, && (11)\\
& \sum_{i \in \mathcal{N}_{\mathrm{P}}} y_{T,i} = 0, && (12)\\
& \sum_{k \in \mathcal{M}} x_{t,k,i,j} - \sum_{k \in \mathcal{M}} x_{t,k,j,i} = 0, \quad t \in \mathcal{T}^{+}, \forall i,j \in \mathcal{N}_{\mathrm{P}}, && (13)\\
& \tau_i \leq \tau_j, \quad \forall i,j \in \mathcal{N}_{\mathrm{P}}, \ i < j. && (14)
\end{aligned}$$

The objective function (3) is the minimization of average download time among peers, $\overline{\tau}$, which is given by

$$\overline{\tau} = \frac{1}{N_{\mathrm{P}}} \sum_{i \in \mathcal{N}_{\mathrm{P}}} \tau_i. \tag{15}$$

Here τ_i $(i \in \mathcal{N}_{\mathrm{P}})$ denotes peer i's download time, which is equal to the length of a time interval during which peer i is a leecher:

$$\tau_i = \sum_{t=0}^{T} y_{t,i}.$$

(4) through (12) represent fundamental constraints of P2P file distribution. In (4) and (5), the server has all pieces at $t = 0$ and each peer has no pieces at $t = 0$, respectively. (6) and (7) describes that self-loop is prohibited and each node can send pieces to others at each time step. (8) means that at time step t, each node can transfer only pieces possessed at $t - 1$. (9) indicates that at time step t, each node requires only pieces that were not possessed at $t - 1$. The upload capacity constraint is given by (10). (11) implies that each node does not require to retrieve the same piece from more than one node. (12) guarantees all peers finish file downloading at least $t = T$.

(13) represents the TFT strategy, under which the numbers of pieces that any pair of leechers exchange should be equal. Note that the left hand side of

(13) denotes the difference of the numbers of pieces exchanged between peer i and peer j at time step t. If both peers i and j are leechers, the left hand side should be zero. Otherwise, at least one of peers i and j is a seed. In this case, the left hand side of (13) also becomes zero because seeds do not require any piece from others as in (9). This situation can be regarded as the seed departure scenario where all peers leave the system immediately after file downloading [5]. In actual P2P file distribution systems, many seeds tend to stay in the system during a relatively short time [13]. The seed departure scenario is the severest scenario among all possible scenarios for peer behavior after file downloading. Note that the seed sojourn scenario, where all seeds permanently remain the system, can also be considered as in [5].

(14) can be regarded as server's policy on the order in which peers finish file downloading. In this paper, we assume that the file downloading should be finished in descending order of peers' upload capacity. We can also consider other kind of policy, e.g., accounting-based prioritization where the server gives preference to accounting peers over others.

In the above formulation, all the variables and constraints except $y_{t,i}$ are linear and $y_{t,i}$ can also be linearized, as shown in A. As a result, problem $\mathrm{P_F}$ can be formulated as ILP.

3.2.2 Streaming distribution case

According to [16], we can also formulate the determination problem of optimal piece flow in the TFT-based P2P streaming distribution, $\mathrm{P_S}$, by modifying problem $\mathrm{P_F}$ as follows.

At first, objective function (3) is replaced with

$$\min \quad \overline{w}, \tag{16}$$

which is the minimization of average play-out delay among peers, $\overline{w}$:

$$\overline{w} = \frac{1}{N_{\mathrm{P}}} \sum_{i \in \mathcal{N}_{\mathrm{P}}} w_i.$$

Here w_i $(i \in \mathcal{N}_{\mathrm{P}})$ denotes peer i's play-out delay,

$$w_i = \max_{k \in \mathcal{M}} w_{i,k}, \quad \forall i \in \mathcal{N}_{\mathrm{P}}, \tag{17}$$

where $w_{i,k} = \sum_{t=0}^{T} (1 - z_{t,k,i}) - (k-1)$ is the waiting time for playing out piece k. Note that $\sum_{t=0}^{T} (1 - z_{t,k,i})$ represents the time to finish retrieving piece k and $k-1$ is the time at which peer i can play out piece k if it has that piece at $t = 0$. As a result, peer i can smoothly play out all pieces if it accepts the initial play-out delay of w_i. In other words, w_i can also be regarded as the maximum jitter that peer i experiences if it starts streaming without any initial play-out delay.

As for the constraints, (4) through (13) should also be kept in the TFT-based P2P streaming distribution because they are the nature of TFT-based

P2P content distribution. (14), which is the constraint of file downloading order, is replaced with the following constraint of play-out delay order:

$$w_i \leq w_j, \quad \forall i, j \in \mathcal{N}_{\mathrm{P}},\ i < j. \tag{18}$$

Since (16) through (18) are linear, $\mathrm{P_S}$ is also ILP.

3.3 Optimal piece flow with topological constraints

$\mathrm{P_F}$ and $\mathrm{P_S}$ in Section 3.2 implicitly assume that the P2P network topology is a full-mesh network, where all peers can select all other peers to obtain pieces under the TFT constraint (13). In acutal systems, e.g., BitTorrent, each peer knows only part of other peers by requesting of a tracker, which is a server that manages the information about all peers in the system.

In terms of autonomous and distributed control, it is desirable that the P2P network topology is as sparse as possible. In other words, it is meaningful for us to find out an optimal topology, which can achieve the minimum average download time with the minimum number of links. We can formulate finding the optimal topology as ILP $\mathrm{P_F^G}$ (resp. $\mathrm{P_S^G}$) by modifying $\mathrm{P_F}$ (resp. $\mathrm{P_S}$) as follows. We first define decision variables $a_{i,j}$ $(i, j \in \mathcal{N}, i < j)$ as

$$a_{i,j} = \begin{cases} 1, & \text{if node } i \text{ and node } j \text{ has a connection,} \\ 0, & \text{otherwise,} \end{cases}$$

where we assume that the connection is undirected. We newly add the following constraints on network topology:

$$a_{i,j} = 1 \quad \forall i \in \mathcal{N}_{\mathrm{D}}, \forall j \in \mathcal{N}_{\mathrm{P}}, \tag{19}$$

$$a_{i,j} = \{0, 1\} \quad \forall i, j \in \mathcal{N}_{\mathrm{P}}, \tag{20}$$

$$x_{t,k,i,j} \leq a_{i,j} \quad \forall t \in \mathcal{T}^{+}, \forall k \in \mathcal{M}, \forall i, j \in \mathcal{N}. \tag{21}$$

(19) indicates that the server can communicate with all peers. Each pair of peers can have a connection as in (20). (21) allows node i to send pieces to node j if there is a connection among them.

By solving $\mathrm{P_F}$ (resp. $\mathrm{P_S}$), we can obtain the optimal average download time $\overline{\tau}^*$ (resp. play-out delay $\overline{w}^*$) when there is no topological constraint. In $\mathrm{P_F^G}$ and $\mathrm{P_S^G}$, we need the following constraint of achieving the optimal objective value,

$$\overline{\tau} = \overline{\tau}^*, \tag{22}$$

$$\overline{w} = \overline{w}^*, \tag{23}$$

respectively.

Finally, the objective function is minimizing the number of connections in the P2P network:

$$\min \sum_{i,j \in \mathcal{N}_{\mathrm{P}}, i<j} a_{ij}. \tag{24}$$

$\mathrm{P_F^G}$ and $\mathrm{P_S^G}$ are also ILP.

4 Analysis of topological influence on optimality of TFT-based P2P content distribution

If the system scale is relatively small, $\mathrm{P_F}$, $\mathrm{P_F^G}$, $\mathrm{P_S}$, and $\mathrm{P_S^G}$ can be directly solved by the existing solver, e.g., CPLEX [1]. In this section, we examine the structure of optimal topology from some numerical results. According to [5,16], we use the following scenario: $\mathcal{N}_\mathrm{D} = \{0\}, \mathcal{N}_\mathrm{P} = \{1, 2, \dots, 6\}$, $C_1 = C_2 = 3, C_3 = C_4 = 2, C_5 = C_6 = 1$. In this scenario, the server's capacity is bottleneck, i.e., $C_0 = 2 < N_\mathrm{P}$. In what follows, we define high-speed peers ($C_i = 3$), middle-speed peers ($C_i = 2$), and low-speed peers ($C_i = 1$). At $t = 0$, server 0 only has all ten pieces ($M = 10$) and all peers start to retrieve the file at $t = 1$. As in [5,16], $\overline{\tau}^*$ and $\overline{w}^*$ become 6.5 and 2.7, respectively.

4.1 Importance of optimal topology

Since the maximum number of links between peers is $N_\mathrm{P}(N_\mathrm{P} - 1)/2$, the number of patterns of the P2P networks becomes $2^{N_\mathrm{P}(N_\mathrm{P}-1)/2}$. If N_P is relatively small, e.g., $N_\mathrm{P} = 6$, we can evaluate all possible topologies. We first examine how the network topology affects average download time $\overline{\tau}$ (resp. average play-out delay $\overline{w}$) in the P2P file distribution (resp. P2P streaming distribution). For this purpose, we calculate optimal average download time $\overline{\tau}^*(g)$ (resp. optimal average play-out delay $\overline{w}^*(g)$) over each possible topology g by solving modified $\mathrm{P_F^G}$ (resp. $\mathrm{P_S^G}$) where (20) is fixed by the corresponding topology g, (22) (resp. (23)) is removed, and (24) is replaced with (3) (resp. (16)).

Fig. 1 illustrates the histogram of $\overline{\tau}^*(g)$ for all possible topologies, where $\overline{\tau}^*(g)$ ranges $[6.5, 14.7]$. Note that all bins except the last bin have size of 0.2 and are half-open (left-closed and right-open) while the last bin ranges $[11.3, 14.7]$. We observe that $\overline{\tau}^*(g)$ is widely distributed, where the fraction of optimal topologies with $\overline{\tau}^*(g) = \overline{\tau} = 6.5$ is less than 2% among all possible topologies. On the contrary, Fig. 2 presents the histogram of $\overline{w}(g)$ for all possible topologies, where $\overline{w}(g)$ ranges $[2.7, 15.8]$. Note that all bins except the last bin have size of 0.2 and are half-open (left-closed and right-open) while the last bin ranges $[7.4, 15.8]$. Comparing with Fig. 1, we observe that the distribution is biased and the fraction of optimal topologies with $\overline{w}(g) = \overline{w} = 2.7$ is more than 0.6.

4.2 Relationship between network density and optimality of distribution

To clarify the relationship between network density, i.e., the total number E of links of topology g, and $\overline{\tau}^*(g)$, we show the minimum, average, and maximum of $\overline{\tau}(g)$ for each E in Fig. 3. As we expect, all of these statistics decrease with E. We observe that $E = 9$ is the minimum value that has some optimal topologies g with $\overline{\tau}(g) = \overline{\tau}$. In Section 4.4, we will confirm that each peer does not necessarily connects, i.e., exchanges pieces, with all other peers to achieve

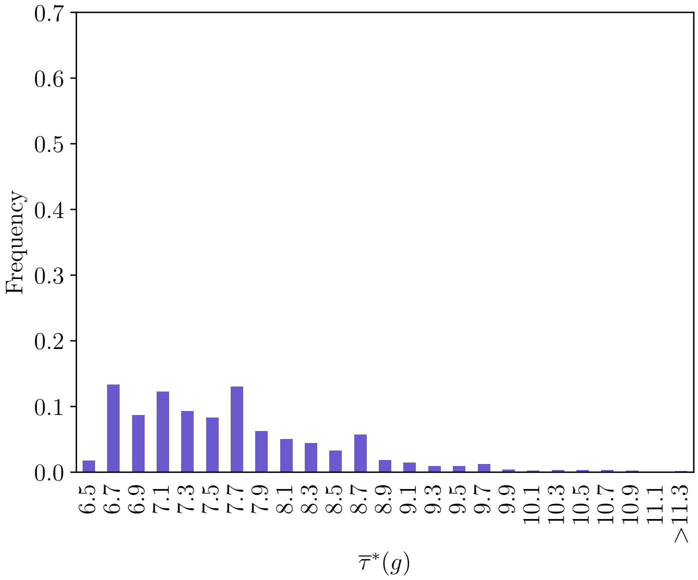


Fig. 1: Histogram of $\overline{\tau}^*(g)$ for all possible topologies.

the optimal piece flow. We also find that the max tends to be large when E is small, which indicates that the topological structure becomes more important for small E.

Fig. 4 illustrates the minimum, average, and maximum of $\overline{w}(g)$ for each E. As in Fig. 3, we observe all of these statistics decrease with E. $E = 6$ is the minimum number of links such that the optimal topologies with $\overline{w}(g) = \overline{w}$ exist. We also find that the range between min and max steeply decrease when $E = 7$, which indicates that the optimality of TFT-based streaming distribution is less sensitive to the topological structure compared with that of TFT-based file distribution.

4.3 Relationship between network size and optimality of distribution

If the network can be divided into multiple sub-networks, i.e., clusters, it will be helpful for the server to maintain the network against dynamic peer arrivals and departures, i.e., peer churn. Fig. 5 illustrates the cluster structure where the whole network is divided into R clusters and server 0 provides C_0/R upload capacity to each cluster.

However, such clustering may increase the average download time and play-out delay. To reveal the degree of this problem, Table 2 demonstrates the relationship between the number N_{P} of peers and the optimal average download time $\overline{\tau}^*$ (play-out delay $\overline{w}^*$) when there is no-cluster structure. Note that the upload capacity C_0 of server is proportion to the network size, i.e., $C_0 = N_{\mathrm{P}}/3$,

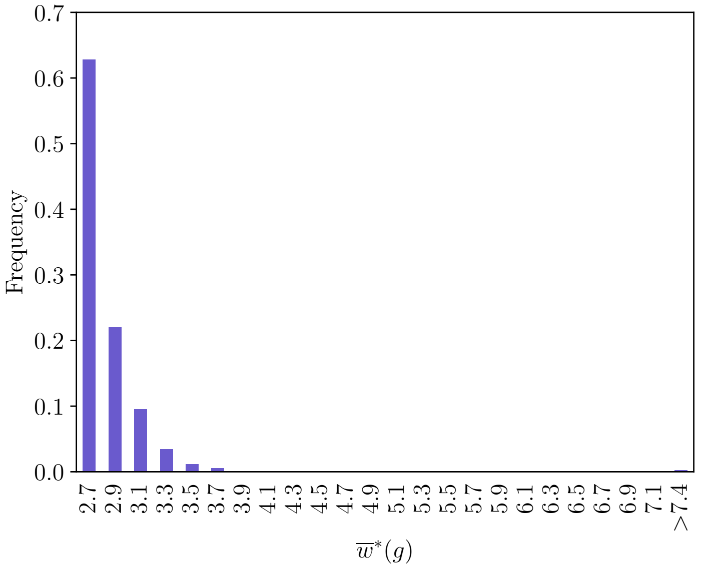


Fig. 2: Histogram of $\overline{w}^*(g)$ for all possible topologies.

Table 2: Impact of clustering structure on system performance.

	$N_\mathrm{P} = 6$ $(C_\mathrm{S} = 2)$	$N_\mathrm{P} = 12$ $(C_\mathrm{S} = 4)$	$N_\mathrm{P} = 18$ $(C_\mathrm{S} = 6)$
$\overline{\tau}^*$	6.5	6.3	6.2
$\overline{w}^*$	2.7	2.6	2.6

and each network consists of $N_\mathrm{P}/3$ high-speed peers, $N_\mathrm{P}/3$ middle-speed peers, and $N_\mathrm{P}/3$ low-speed peers. We observe that $\overline{\tau}^*$ and $\overline{w}^*$ do not almost change with N_P, which indicates that we can divide the whole system of $N_\mathrm{P} = 12$ (resp. $N_\mathrm{P} = 18$) into two (resp. three) sub-systems of $N_\mathrm{P} = 6$ while keeping the system performance. In other words, we can estimate the behavior of larger systems by regarding the whole system consists of multiple small-scale sub-systems even if it is difficult to directly analyze the whole system, due to the computation complexity.

4.4 Structure of optimal topology

Next, we reveal the structure of optimal topology through the numerical results. Fig. 7 illustrates all the six optimal topologies with $\overline{\tau}(g) = \overline{\tau} = 6.5$ and the minimum number of links, i.e., $E = 9$, for TFT-based P2P file distribution. Since peers with identical upload capacity are interchangeable, e.g., peer 1 and peer 2, we find Figs. 6a and 6b (resp. Figs. 6c through 6f) are equivalent. As a result, there are only two optimal topologies.

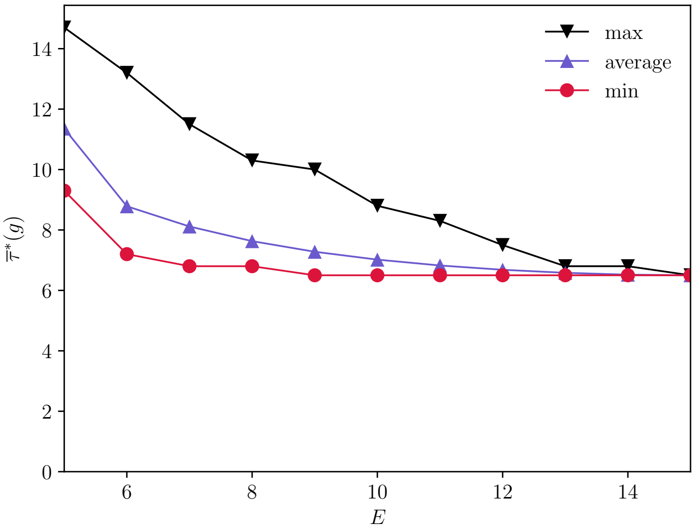


Fig. 3: min, average, and max of $\overline{\tau}^*(g)$ for each E.

On the other hand, Fig. 7 presents all the ten optimal topologies with $\overline{w}(g) = \overline{w} = 2.7$ and the minimum number of links, i.e., $E = 6$, for TFT-based P2P streaming distribution. We can also find Figs. 7a through 7e (resp. Figs. 7f through 7j) are equivalent by taking account of the interchangeability of peers with the same upload capacity, and thus there are only two optimal topologies.

It is desirable that the topological structure is as simple as possible from the viewpoint of practicality. We observe that Fig. 6a has a simple and interesting structure, i.e., hierarchical and circular structure. Note that the topology in Fig. 6a is a supergraph of the topology in Fig. 7a. For simplicity, we assume that all N_P peers can be equally divided into three groups, each of which consists of $L_\mathrm{P} = N_\mathrm{P}/3$ peers whose upload capacities are given as follows:

$$C_i = \begin{cases} 3 & (i = 1, 2, \ldots, L_\mathrm{P}), \\ 2 & (i = L_\mathrm{P}+1, L_\mathrm{P}+2, \ldots, 2L_\mathrm{P}), \\ 1 & (i = 2L_\mathrm{P}+1, 2L_\mathrm{P}+2, \ldots, N_\mathrm{P}). \end{cases}$$

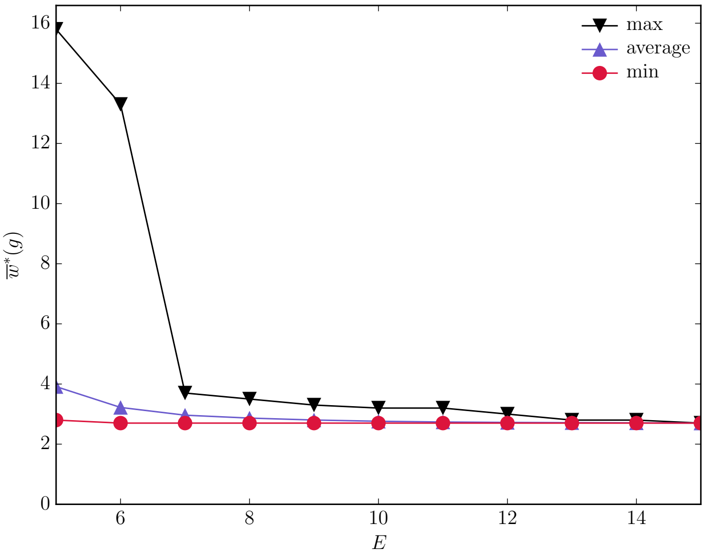


Fig. 4: min, average, and max of $\overline{w}^*(g)$ for each E.

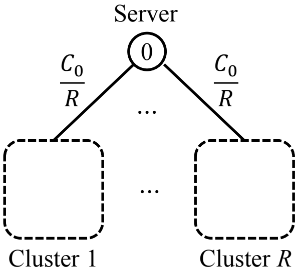
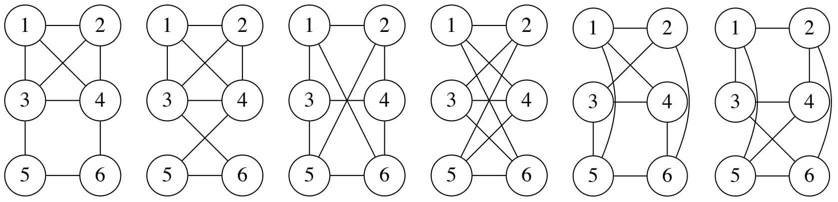


Fig. 5: Clustering structure.

The hierarchical and circular network topology can be generalized as shown in Fig. 8, which has links between peers as follows:

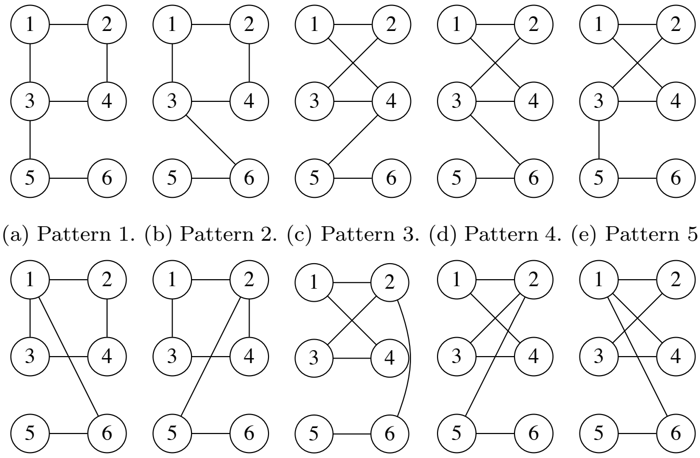
$$a_{i,j} = \begin{cases} 1 & (i = lL_{\mathrm{P}} + 1, lL_{\mathrm{P}} + 2, \ldots, (l+1)L_{\mathrm{P}}, \\ & \quad j = lL_{\mathrm{P}} + (i+1) \bmod L_{\mathrm{P}}, l = 0, 1, 2), \quad (25) \\ 1 & (i = 1, 2, \ldots, 2L_{\mathrm{P}}, j = i + L_{\mathrm{P}}), \quad (26) \\ 1 & (i = 1, 2, \ldots, L_{\mathrm{P}}, \\ & \quad j = L_{\mathrm{P}} + (i+1) \bmod L_{\mathrm{P}}), \quad (27) \\ 0 & \text{otherwise.} \end{cases}$$

The hierarchical and circular network topology has the following properties.



(a) Pattern 1. (b) Pattern 2. (c) Pattern 3. (d) Pattern 4. (e) Pattern 5. (f) Pattern 6.

Fig. 6: Optimal network topology for TFT-based P2P file distribution ($N_{\mathrm{P}} = 6, E = 9$).



(a) Pattern 1. (b) Pattern 2. (c) Pattern 3. (d) Pattern 4. (e) Pattern 5.

(f) Pattern 6. (g) Pattern 7. (h) Pattern 8. (i) Pattern 9. (j)Pattern 10.

Fig. 7: Optimal network topology for TFT-based P2P streaming distribution ($N_{\mathrm{P}} = 6, E = 6$).

- Hierarchical structure based on upload capacity:
  There are three layers of peers: top layer of high-speed peers, middle layer of middle-speed peers, and bottom layer of low-speed peers. Two adjacent layers have links, e.g., e_3, e_5, which are given by (26). In addition, the top layer and middle layer has additional links, e.g., e_2, which are given by (27). There is no link between top layer and bottom layer. Recall that it has been pointed out that the server should distribute pieces indirectly to peers with large upload capacity via those with small upload capacity [5]. Hierarchical structure based on upload capacity is also important.
- Circular structure at each layer: Peers have two links with adjacent peers at the same layer, e.g., e_1, e_4, e_5, which are given by (25).
- Degree depending on upload capacity and hierarchical structure: The number of links, i.e., degree, of peers basically depends on their upload capacity. In addition, middle-speed peers at the middle layer have more links com-

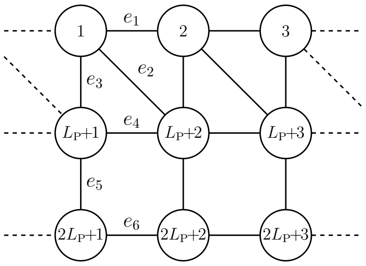
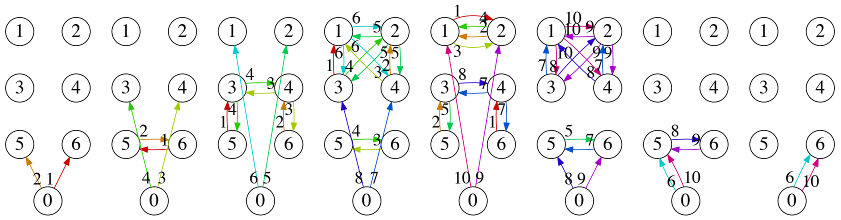


Fig. 8: Hierarchical and circular network topology.



(a) $t = 1$. (b) $t = 2$. (c) $t = 3$. (d) $t = 4$. (e) $t = 5$. (f) $t = 6$. (g) $t = 7$. (h) $t = 8$.

Fig. 9: Optimal piece flow per time for TFT-based P2P file distribution (optimal topology, $N_\mathrm{P} = 6, E = 9$).

pared with peers at other layers because they play roles of bridging top layer and bottom layer.

We confirm that the hierarchical and circular network topology is optimal in case of $N_\mathrm{P} = 9, 12, 18$.

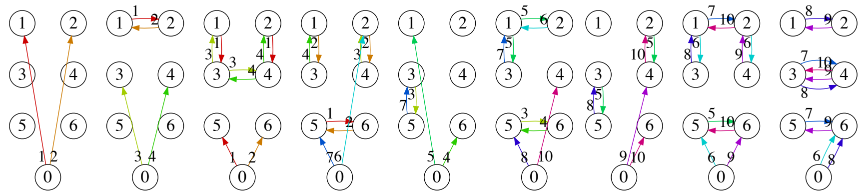
Finally, we derive the number of links in the hierarchical and circular network topology. As shown in Fig. 8, we observe that the structure composed of nodes $1, L_\mathrm{P} + 1, 2L_\mathrm{P} + 1$ and links $e_1, e_2, \ldots, e_6$ is repeated. Recall each layer has $L_\mathrm{P} = N_\mathrm{P}/3$ peers. As a result, the number $E^*(N_\mathrm{P})$ of links in the hierarchical and circular topology is given as follows:

$$E^*(N_\mathrm{P}) = \begin{cases} 2N_\mathrm{P} - 3 & (N_\mathrm{P} = 6), \\ 2N_\mathrm{P} & (N_\mathrm{P} = 9, 12, \ldots). \end{cases}$$

$E^*(N_\mathrm{P})$ is a linear function of N_P and much smaller than those of full-mesh network, i.e., $N_\mathrm{P}(N_\mathrm{P} - 1)/2$ links. Since each peer has a constant number of links, the hierarchical and circular topology has highly autonomous and distributed property.

4.5 Structure of optimal piece flow

Finally, we confirm the optimal piece flow over the optimal topology. Fig. 9 depicts the optimal piece flow of TFT-based P2P file distribution under the



(a) $t = 1$. (b) $t = 2$. (c) $t = 3$. (d) $t = 4$. (e) $t = 5$. (f) $t = 6$. (g) $t = 7$. (h) $t = 8$. (i) $t = 9$.

Fig. 10: Optimal piece flow per time for TFT-based P2P streaming distribution (optimal topology, $N_\mathrm{P} = 6, E = 6$).

optimal topology in Fig. 6a, i.e., $N_\mathrm{P} = 6$ and $E = 9$. Each circle indicates a node with its node ID $i \in \mathcal{N}$ and each arrow with piece ID $k \in \mathcal{M}$ shows the flow of piece k from node $i \in \mathcal{N}$ to peer $j \in \mathcal{N}_\mathrm{P}$. As mentioned in [5], in case of file distribution under a full-mesh network, the optimal piece flow requires that the server sends pieces to peers in ascending order of their upload capacities at the early stage of distribution. We can confirm the same tendency is kept in optimal piece flow under the optimal topology as in Fig. 9. We also observe that these pieces provided by the server can be locally exchanged among connected peers.

Fig. 10 illustrates the optimal piece flow of TFT-based P2P streaming distribution under the optimal topology in Fig. 10a, i.e., $N_\mathrm{P} = 6$ and $E = 6$. In case of the optimal piece flow of streaming distribution under a full-mesh network, the server sends pieces to peers in descending order of their upload capacities at the early stage of distribution [16]. This characteristic also appears in the optimal piece flow under the optimal topology. We also confirm that neighboring peers can effectively exchange the pieces retrieved from the server as in the file distribution case.

5 Conclusions

In this paper, we investigated the topological influence on the optimality of TFT-based P2P content distribution. We first formulated the topology optimization problem as ILP, which finds an optimal topology where the optimal piece flow is achieved with the minimum number of links. Through numerical results, we first showed the topological structure becomes more important in case of file distribution. We also revealed that both file and streaming distribution can achieve the optimal piece flow under the hierarchical and circular topology with a smaller number of links, i.e., $O(N_\mathrm{P})$, compared with that of full-mesh network, i.e., $O(N_\mathrm{P}^2)$. In addition, we found that the whole network can be divided into multiple sub-networks, i.e., clusters, while keeping the system performance. Finally, we also showed the optimal piece flow over the optimal topology, in which the server's optimal piece sending strategy is the same as that in a full-mesh network and peers can effectively exchange the pieces retrieved from the server with neighboring peers.

A Linearization of products of binary variables

The product of variables are nonlinear but it can be transformed into linear expressions if all variables are binary [3]. In particular,

$$y = x_1 x_2 \cdots x_k, \qquad x_i = \{0,1\}, \quad (i = 1, 2, \ldots, k)$$

can be rewritten to be the following linear expressions:

$$\begin{aligned} &(k-1) - \sum_{i=1}^{k} x_i + y \geq 0, \\ &x_i - y \geq 0, \quad x_i = \{0,1\}, \quad (i = 1, 2, \ldots, k). \end{aligned}$$

With this technique, nonlinear terms in (2) and (13) can be linearized.

References

1. IBM ILOG CPLEX Optimizer. `https://www.ibm.com/products/ilog-cplex-optimization-studio` (Accessed 31 Oct. 2018)
2. BitTorrent Inc.: BitTorrent. `http://www.bittorrent.com/` (Accessed 31 Oct. 2018)
3. Chen, D., Batson, R.G., Dang, Y.: Applied Integer Programming. Wiley (2010)
4. Garlaschelli, D., Loffredo, M.I.: Patterns of Link Reciprocity in Directed Networks. Physical Review Letters **93**(26), 1–4 (2004)
5. Hasegawa, M., Sasabe, M., Takine, T.: Analysis of Optimal Scheduling in Tit-for-Tat-based P2P File Distribution. IEICE Transactions on Communications **E97-B**(12), 2650–2657 (2014)
6. 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)
7. Hughes, D., Coulson, G., Walkerdine, J.: Free Riding on Gnutella Revisited: The Bell Tolls? IEEE Distributed Systems Online **6**(6), 1–18 (2005)
8. Jamjoom, H., Shin, K.G.: Persistent Dropping: An Efficient Control of Traffic Aggregates. In: Proc. of SIGCOMM '03, pp. 287–298 (2003)
9. Li, B., Ma, M., Jin, Z., Zhao, D.: Investigation of a Large-Scale P2P VoD Overlay Network by Measurements. Peer-to-Peer Networking and Applications **5**(4), 398–411 (2012)
10. Marza, V., Dehghan, M., Akbari, B.: A New Peer-to-Peer Topology for Video Streaming Based on Complex Network Theory. Journal of Systems Science and Complexity **28**(1), 16–29 (2015)
11. Microsoft: Windows Update Delivery Optimization. `https://docs.microsoft.com/en-us/windows/deployment/update/waas-delivery-optimization` (Accessed 31 Oct. 2018)
12. Neglia, G., Presti, G.L., Zhang, H., Towsley, D.: A Network Formation Game Approach to Study BitTorrent Tit-for-Tat. In: Proceedings of the 1st EuroFGI International Conference on Network Control and Optimization, pp. 13–22 (2007)
13. Pouwelse, J., Garbacki, P., Epema, D., Sips, H.: The Bittorrent P2P File-Sharing System: Measurements and Analysis. Peer-to-Peer Systems IV pp. 205–216 (2005)
14. PPLive Inc.: PPTV. `http://www.pptv.com/`
15. Rousseau, R., Zhang, L.: Betweenness Centrality and Q-Measures in Directed Valued Networks. Scientometrics **75**(3), 575–590 (2008)

16. Sasabe, M.: Analysis of Optimal Piece Flow in Tit-for-Tat-Based P2P Streaming. Computer Networks **139**(5), 60–69 (2018). DOI 10.1016/j.comnet.2018.04.004
17. Stutzbach, D., Rejaie, R., Sen, S.: Characterizing Unstructured Overlay Topologies in Modern P2P File-Sharing Systems. IEEE/ACM Transactions on Networking **16**(2), 267–280 (2008)
18. Su, M., Zhang, H., Du, X., Fang, B., Guizani, M.: A Measurement Study on the Topologies of BitTorrent Networks. IEEE Journal on Selected Areas in Communications **31**(9), 338–347 (2013)
19. Vu, L., Gupta, I., Nahrstedt, K., Liang, J.: Understanding Overlay Characteristics of a Large-scale Peer-to-peer IPTV System. ACM Trans. Multimedia Comput. Commun. Appl. **6**(4), 31:1–31:24 (2010)
20. Wu, C., Li, B., Zhao, S.: Exploring Large-Scale Peer-to-Peer Live Streaming Topologies. ACM Transactions on Multimedia Computing, Communications, and Applications **4**(3), 1–23 (2008)
21. Yin, B., Lu, X., Huang, J., Kang, Y.: Analysis of Topology Dynamics for Unstructured P2P Networks. Computer Communications **80**, 72–81 (2016)
22. Zhang, H., Neglia, G., Towsley, D., Presti, G.L.: On Unstructured File Sharing Networks. In: Proceedings of 26th IEEE International Conference on Computer Communications (INFOCOM '07), pp. 2189–2197 (2007)
23. Zhong, L., Kihl, M., Wang, X.: Topological Model and Analysis of the P2P BitTorrent Protocol. International Journal of System Control and Information Processing **1**(1), 54–70 (2012)