P2P Video-on-Demand: Steady State and Scalability

Niklas Raatikainen, Sasu Tarkoma, Petri Savolainen
Helsinki Institute for Information Technology
Email: firstname.lastname@hiit.fi

Samuli Aalto, Pasi Lassila
TKK Helsinki University of Technology
Email: firstname.lastname@tkk.fi

Abstract—**The fundamental P2P principle that downloading peers help other peers can be applied in the context of on-demand streaming of stored content (video-on-demand). This represents a demanding application combining aspects of other well-known P2P applications, i.e., live streaming and traditional file sharing. We seek to determine the conditions when the system operates properly. A deterministic fluid model is derived that explicitly takes into account the video transfer and playback phases. The analytical results comprise the analysis of the uniqueness of the steady state and its local stability. These results are complemented with extensive simulations from the corresponding stochastic model, as well as traces from a more realistic BitTorrent simulator implementing a windowing algorithm. Finally, we provide a synthesis of the analytical results and the numerical validations to characterize the prerequisites under which the piece retrieval rate exceeds the video viewing rate and under which conditions this is achieved independent of peer arrival rate.**

I. Introduction

Modern peer-to-peer (P2P) content distribution networks such as BitTorrent [1] achieve scalability by utilizing the fundamental idea that the peers not only act as clients downloading content from other peers but also serve the other peers by uploading onwards the downloaded content. While P2P content distribution technologies are widely used for file sharing and live media streaming, they are less common in the closely related, but yet distinct case of *on-demand streaming of stored media* [2], of which a well-known example is *video-on-demand* (VoD). However, the popularity of client-server based VoD services such as YouTube raises the question, whether, and under which conditions, P2P technologies could be used to alleviate the high cost of building and maintaining the infrastructures of such systems.

In P2P VoD, the whole stored media file needs to be retrieved (cf., file sharing) at such a rate which allows, as soon as possible, the pieces to be played back in sequential order at the media playback rate (cf., live streaming). Thus, there are two phases, the *transfer* and the *playback* phases, which are partly overlapping. If the retrieval rate is sufficient, the playback phase extends beyond the transfer phase. Essential aspects affecting the performance of the system include the efficiency of the piece exchange. Another one is the notion of *altruism*, i.e., the peers that have finished the transfer phase become seeds and can help the system perform better by allowing others to download from it. In our work, we make a minimal assumption on the level of altruism required from a peer: an altruistic peer is assumed to allow others to download from it only for as long as it is in the playback phase.

In this paper we present two models, a deterministic fluid model and a stochastic model, of P2P VoD systems. We seek to provide insight on fundamental questions about the performance and scalability of the system. More specifically, our contributions include:

- Presenting a fluid model that allows both download and upload constrained solutions and also explicitly takes into account the transfer and the playback phases, which are not done, e.g., in [2].
- Explicitly characterizing the conditions under which the system has a positive, unique and asymptotically stable steady-state solution.
- Validating the fluid model by comparing it to results from simulating the corresponding stochastic model and to traces from a realistic BitTorrent simulator.
- Utilizing the analytical and the numerical results to give closed-form conditions under which the piece retrieval rate exceeds the video viewing rate (i.e., QoS is acceptable) and under which conditions this is achieved independent of peer arrival rate (i.e., system is scalable).

Notably, our results for example show that to achieve a piece retrieval rate exceeding the video viewing rate, the efficiency parameter is critical, i.e., it must be large enough. On the other hand, the better-performing, download-constrained steady state can always be achieved with a high enough fraction of altruistic seeds or a sufficient number of permanent seeds.

The paper is organized as follows. Related work is discussed in detail in Section II. The models and the analytical results are presented in Section III. A synthesis of the results and interpretation with respect to scalability is given in Section IV. The numerical validation results are found in Section V, and conclusions are given in Section VI.

II. Related work

P2P live streaming has been analyzed recently in many papers, see, e.g., [3] analyzing the impact of various peer/piece selection policies or [4], where a stochastic model is developed for analyzing the probability that peers can download the stream at a sufficiently high rate. However, in this paper our objective is to consider the case of on-demand streaming and to develop fluid models in the spirit of [5].

Qiu and Srikant [5] develop a simple fluid model to analyze the performance of a BitTorrent-like P2P *file sharing* system under a steady flow arrival scenario. Their model, which describes the time-evolution of the system by differential

equations, is an extension of the deterministic counterpart of an earlier Markovian model by Yang and de Veciana [6].

The key parameters of these models are the arrival rate of new peers, λ, the efficiency of P2P file sharing, η, and the departure rate of seeds, γ, which reflects their selfishness aspect. The efficiency parameter η comprises the effect of the piece selection policy, the number of downloading connections, and the number of pieces. Qiu and Srikant conclude that $\eta \approx 1$ whenever the number of pieces is sufficiently high. Both models also assume a homogeneous peer population with joint download and upload rates, c and μ, respectively. In fact, Yang and de Veciana implicitly assume that the system is upload-constrained (i.e., $c \gg \mu$), while Qiu and Srikant allow any positive values of c and μ. Qiu and Srikant have also an additional parameter θ modeling the rate at which leechers abort the file transfer, but Yang and de Veciana do not consider such events (i.e., $\theta = 0$).

Using the notation given above and assuming that $\theta = 0$, the fluid model by Qiu and Srikant [5] can be expressed as:

$$\begin{cases} x'(t) = \lambda - \min\{cx(t), \mu(\eta x(t) + y(t))\}, \\ y'(t) = \min\{cx(t), \mu(\eta x(t) + y(t))\} - \gamma y(t), \end{cases} \quad (1)$$

where the two state variables are the number of leechers, $x(t)$, and the number of seeds, $y(t)$. The global stability of this fluid model (for $\eta > 0$) is shown in [7]. Extension of [5] to heterogeneous users has been given in [8], [9].

Inspired by Qiu and Srikant [5], Parvez et al. [2] develop fluid models to analyze the performance of a BitTorrent-like P2P *on-demand streaming*. They start from the premise that the system is upload-constrained leading to the following fluid model for the rarest-first piece selection policy:

$$\begin{cases} x'(t) = \lambda - \mu(\eta x(t) + y(t)), \\ y'(t) = \mu(\eta x(t) + y(t)) - \gamma y(t). \end{cases} \quad (2)$$

Like Qiu and Srikant, Parvez et al. conclude that $\eta \approx 1$ at least for "most scenarios of interest". Their key point, however, is that the fluid model (2) can also be used to model a smart variant of the in-order piece selection policy, which is superior to the rarest-first policy in the on-demand streaming case.

While the assumption that $\eta \approx 1$ is plausible for file sharing systems with the rarest-first policy, it may not be the case for on-demand-streaming. For example, simulations with realistic P2P VoD systems in [10] have shown that the efficiency of the piece exchange is limited by the used windowing mechanism. In fact, for live streaming Tewari and Kleinrock [11] argue that $\eta \approx 1 - 1/N$, where N is the window size which may be small depending on the real-time requirements. As on-demand-streaming systems typically employ an adaptive window size allowing for more flexibility in the piece selection than strict live streaming, the achievable efficiency of on-demand-streaming lies somewhere between live streaming and file sharing.

Our fluid model for an on-demand streaming system is also inspired by Qiu and Srikant [5]. However, the essential difference between our model and that of [2] is in the handling of the playback phase and the modeling of selfishness. In fact, model (2) totally ignores the playback phase and simply assumes that any seed, whether it has played back the media file or not, departs with rate γ. Instead of that, we include the playback phase explicitly in our model and make an assumption about playback phase altruism: a fraction of peers, i.e., the *altruistic* peers, will stay in the system as seeds until the end of the playback phase and then leave (which is a worst case scenario), while the selfish peers will leave the system the moment they are finished with their transfer phase. Nor do we restrict ourselves to the upload-constrained case but allow c and μ to take any positive values. We generalize the previous models by adding a finite number of original seeds that stay in the system permanently. In addition to a fluid model, we also develop a stochastic counterpart that is used for numerical purposes to get a more accurate view of the system, as well as to work as a bridge between the fluid model and windowed BitTorrent simulation results presented in [10].

III. SYSTEM MODELING AND ANALYSIS

Consider a P2P VoD application. From here on we will use BitTorrent terminology, but the fluid and stochastic model are both applicable to any P2P application based on dividing a file into pieces and the mutual exchange of those pieces between the users. Our purpose is to describe the system dynamics related to the sharing of a single video file assuming a homogeneous peer population and a steady flow arrival scenario. Indeed, it has been shown in [12] through extensive analysis of YouTube usage that the demand for new videos is reasonably predictable (constant) on the order of days.

A. Modeling framework

Let m denote the size of the video file to be shared (in bits). The video is played back at a constant rate w (bits per time unit). New peers arrive at rate λ (arrivals per time unit on average). Each peer is connected to the network over an asymmetric access link (e.g., an ADSL link) with a download capacity of d (bits per time unit) and upload capacity u (bits per time unit). The corresponding download and upload rates are: $c = d/m$ and $\mu = u/m$ (file transfers per time unit).

The life span of a peer consists of two phases, the file *transfer phase* and the *playback phase*, which are partly overlapping. The first part of the transfer phase (called the startup delay) lasts until there are sufficiently pieces in the playback buffer to start the playback phase. After that, the video transfer and playback proceed in a parallel manner until the entire video is transferred. It is desirable that the transfer rate is greater than the playback rate so that the video can be played back without any breaks or delays. In such a case, the playback phase extends beyond the transfer phase.

We assume that the startup delay is negligible. This is possible to achieve with a sufficiently small window size. In addition, we assume that

$$d > w, \quad (3)$$

which is a necessary requirement for a VoD system in order that the video can be played back with good quality, since the

transfer rate for a downloading peer is always upper bounded by its own access link rate.

During the transfer phase, the peer is called a *leecher*. According to the fundamental P2P principle, leechers help each other. Let $\eta \in [0,1]$ denote the efficiency of this operation, i.e., what fraction of peer upload capacity is used to upload to other peers. As for the file sharing application, the efficiency parameter η comprises the total effect of the P2P protocol in use. In the case of windowed BitTorrent, parameters that affect η are the basic efficiency of BitTorrent, the size of the window used, the number of downloading connections, and the number of pieces. A more detailed examination of efficiency in the windowed BitTorrrent case can be found in [10].

An altruistic leecher becomes a *seed* as soon as its own video file download is completed. Let $\zeta \in [0,1]$ denote the fraction of altruistic peers, which continue to upload to leechers even after their own transfer phase. Non-altruistic peers are assumed to leave the system immediately after the transfer phase.

We assume that each altruistic peer stays in the system (exactly) until the end of the playback phase. If the video is played back without any breaks or delays, the length of the playback phase equals $z = m/w$. It follows from (3) that

$$z > 1/c. \qquad (4)$$

In addition to the non-permanent seeds, we allow a number of original seeds, say k, that stay in the system permanently.

B. Fluid model

Let $x(t)$ and $y(t)$ denote the number of leechers and seeds, respectively, at time t. We first derive a deterministic fluid model for characterizing the dynamics of $x(t)$ and $y(t)$.

The number of leechers, $x(t)$, increases with rate λ. On the other hand, it is decreased with the rate at which file transfers complete. Assuming that there are no bottlenecks in the core network, this completion rate $\phi(t)$ is determined by the minimum of the total download rate and the total upload rate (cf. [5], [2]),

$$\phi(t) = \min\{cx(t), \mu(\eta x(t) + y(t) + k)\}. \qquad (5)$$

Thus, we have $x'(t) = \lambda - \phi(t)$.

As mentioned above, an altruistic leecher becomes a seed after the transfer phase. Thus, the number of seeds, $y(t)$, increases with rate $\phi(t)\zeta$. Characterizing the rate at which seeds depart from the system is more difficult and requires some approximations given below.

We base our model on *initially* making the optimistic assumption that the transfer rate is sufficient so that the video can be played back without any breaks or delays. Since we also assumed that the startup delay is negligible and altruistic peers leave the system as soon as they complete the playback phase, this fixes the total time that such a peer spends in the system to be equal to z. Now we make an *approximative* assumption that the system is quasi-stationary at every point in time t. Then by Little's result, the average time that an altruistic peer stays as a seed equals $z - x(t)/\lambda$. Thus, the total

departure rate of seeds becomes $y(t)/(z - x(t)/\lambda)$ implying that $y'(t) = \zeta\phi(t) - y(t)/(z - x(t)/\lambda)$.

All in all, our fluid model is as follows:

$$\begin{cases} x'(t) = \lambda - \phi(t), \\ y'(t) = \zeta\phi(t) - \frac{y(t)}{z - x(t)/\lambda}, \end{cases} \qquad (6)$$

where $\phi(t)$ is given in (5). The model belongs to the class of so called switched nonlinear systems due to the presence of the $\min$-function, see [13].

Note that there are no a priori guarantees that the difference $z - x(t)/\lambda$ stays positive. If this is not the case, the differential equation system behaves in an unstable manner, which can be interpreted as a sign of problems in the playback quality of the P2P VoD application.

C. Steady-state analysis

Model (6) allows us to explicitly solve the equilibrium of the system by setting $x'(t) = y'(t) = 0$ and solving for the corresponding values of x and y. However, due to the $\min$-operation in (5), we need to separately consider whether the system is download or upload constrained at the equilibrium, which results in two different equilibria.

Let us denote by $\bar{x}$ and $\bar{y}$ the equilibrium values of (6). The system is *download-constrained* at the equilibrium if $c\bar{x} \leq \mu(\eta\bar{x} + \bar{y} + k)$, implying that

$$\begin{cases} \bar{x}_{\mathrm{d}} &=& \frac{\lambda}{c}, \\ \bar{y}_{\mathrm{d}} &=& \zeta\lambda(z - \frac{1}{c}), \end{cases} \qquad (7)$$

where the subscript (d) refers to the download-constrained solution. Now the constraint $c\bar{x}_{\mathrm{d}} \leq \mu(\eta\bar{x}_{\mathrm{d}} + \bar{y}_{\mathrm{d}} + k)$ is equivalent to

$$\frac{1}{\mu} \leq \frac{\eta}{c} + \zeta(z - \frac{1}{c}) + \frac{k}{\lambda}, \qquad (8)$$

which implies that $\frac{1}{\mu} - \frac{k}{\lambda} < z$ is a necessary condition and $\frac{1}{\mu} - \frac{k}{\lambda} < 0$ a sufficient condition for the existence of a download-constrained solution. The area limited by (8) is illustrated in Figure 1 for varying η and ζ.

We also see that $\bar{x}_{\mathrm{d}} > 0$ for sure and $\bar{y}_{\mathrm{d}} > 0$ by (4) so that the solutions are meaningful in the download-constrained case. Furthermore, (4) says that the equilibrium transfer rate for a leecher exceeds the watching rate in this case.

On the other hand, if $c\bar{x} > \mu(\eta\bar{x} + \bar{y} + k)$, then the system is *upload-constrained* at the equilibrium and we have

$$\begin{cases} \bar{x}_{\mathrm{u}} &=& \frac{\lambda}{\eta - \zeta}(\frac{1}{\mu} - \zeta z - \frac{k}{\lambda}), \\ \bar{y}_{\mathrm{u}} &=& \frac{\zeta\lambda}{\eta - \zeta}(-\frac{1}{\mu} + \eta z + \frac{k}{\lambda}), \end{cases} \qquad (9)$$

which requires (to keep $\bar{x}_{\mathrm{u}}$ and $\bar{y}_{\mathrm{u}}$ finite) that

$$\eta \neq \zeta. \qquad (10)$$

In this case, the constraint $c\bar{x}_{\mathrm{u}} > \mu(\eta\bar{x}_{\mathrm{u}} + \bar{y}_{\mathrm{u}} + k)$ is equivalent to

$$\begin{cases} \frac{1}{\mu} > \frac{\eta}{c} + \zeta(z - \frac{1}{c}) + \frac{k}{\lambda}, & \text{if } \eta > \zeta, \\ \frac{1}{\mu} < \frac{\eta}{c} + \zeta(z - \frac{1}{c}) + \frac{k}{\lambda}, & \text{if } \eta < \zeta. \end{cases} \qquad (11)$$

The area limited by (11) is illustrated in Figure 1 for varying η and ζ.

Additionally, for the solution to be meaningful, we require that $\bar{x}_u > 0$ and $\bar{y}_u > 0$. The former one follows from (11), but the latter one leads to the following additional constraints:

$$\begin{cases} \zeta < \frac{1}{z}(\frac{1}{\mu} - \frac{k}{\lambda}) < \eta, & \text{if } \eta > \zeta, \\ \eta < \frac{1}{z}(\frac{1}{\mu} - \frac{k}{\lambda}) < \zeta, & \text{if } \eta < \zeta, \end{cases} \quad (12)$$

which implies that $0 < \frac{1}{\mu} - \frac{k}{\lambda} < z$ is a necessary condition for the existence of a positive upload-constrained solution. It is also straightforward to check that the equilibrium transfer rate for a leecher exceeds the watching rate in this upload-constrained case if and only if $\bar{y}_u > 0$.
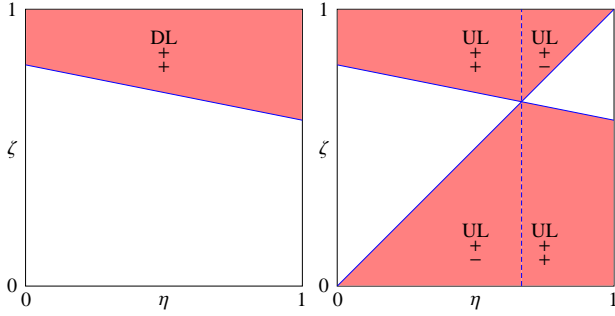

Fig. 1. *Left panel*: Download-constrained solution with $+/+$ expressing that $\bar{x}_d > 0$ and $\bar{y}_d > 0$ in this area. *Right panel*: Upload-constrained solution with $+/+$ $[+/-]$ expressing that $\bar{x}_u > 0$ and $\bar{y}_u > 0$ $[\bar{y}_u < 0]$ in this area.

To summarize the existence of positive equilibrium solutions for the fluid model (6):

 (i) If $\frac{1}{\mu} - \frac{k}{\lambda} \geq z$, then there are no positive solutions in the (η, ζ)-box.
 (ii) If $0 < \frac{1}{\mu} - \frac{k}{\lambda} < z$, then there are four solution areas: the upper-left area, where there are two different positive solutions (DL and UL), the upper-right area, where there is a unique download-constrained solution (DL), the lower-right area, where there is a unique upload-constrained solution (UL), and the lower-left area, where there are no positive solutions.
 (iii) If $\frac{1}{\mu} - \frac{k}{\lambda} \leq 0$, then there is a unique download-constrained solution.

In case (ii), which is illustrated in Figure 2, the joint corner point of the four solution areas satisfies

$$(\eta, \zeta) = (\frac{1}{z}(\frac{1}{\mu} - \frac{k}{\lambda}), \frac{1}{z}(\frac{1}{\mu} - \frac{k}{\lambda})),$$

and it is sliding along the diagonal from $(0,0)$ to $(1,1)$ as λ increases or z, μ, or k decreases.

Finally we note that at the equilibrium $(\bar{x}, \bar{y})$, by Little's result we must have

$$\frac{\bar{x}}{\lambda} + \frac{\bar{y}}{\zeta\lambda} = z.$$

It can be easily verified that the solutions (7) and (9) indeed both satisfy the above relation.
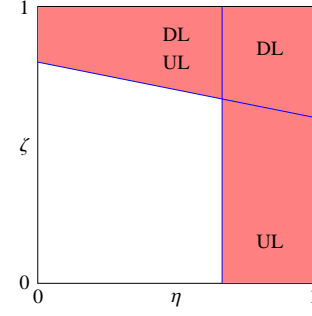

Fig. 2. Solution areas, where DL [UL] refers to a positive download [upload] constrained solution. The horizontal bordering line satisfies $\frac{1}{\mu} = \frac{\eta}{c} + \zeta(z - \frac{1}{c}) + \frac{k}{\lambda}$ and the vertical bordering line satisfies $\eta = \frac{1}{z}(\frac{1}{\mu} - \frac{k}{\lambda})$.

D. Local stability analysis

We can apply the methods from control theory to analyze the stability of (6). Analyzing the global stability of the model is difficult due to the nonlinear and switched nature of the model, see [13]. Instead we focus on the local stability of the system. In particular, our approach is to linearize the system of differential equations around its equilibrium, and apply the control theory of linear systems around the equilibrium. Thus, the stability results relate to the local stability of the original system (6) around the equilibrium. Note that local stability of the equilibrium does not, in general, guarantee that the system converges to the equilibrium from arbitrary initial conditions.

In general, our differential equations can be written as

$$\begin{cases} x'(t) &= f_1(x, y), \\ y'(t) &= f_2(x, y). \end{cases} \quad (13)$$

Recall that the equilibrium is denoted by $(\bar{x}, \bar{y})$ and is obtained by setting the right hand sides equal to zero in (13).

We can study how the equilibrium $(\bar{x}, \bar{y})$ is reached by linearizing the system near $(\bar{x}, \bar{y})$. The local stability is then characterized by the roots of the eigenvalue equation

$$z^2 - (a_1 + a_4)z + a_1 a_4 - a_2 a_3 = 0, \quad (14)$$

where the constants a_i are given by the elements of the Jacobian matrix evaluated at the equilibrium

$$\begin{aligned} \mathbf{J}(\bar{x}, \bar{y}) &= \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{\partial f_1(x, y)}{\partial x} & \dfrac{\partial f_1(x, y)}{\partial y} \\ \dfrac{\partial f_2(x, y)}{\partial x} & \dfrac{\partial f_2(x, y)}{\partial y} \end{pmatrix} \Bigg|_{x=\bar{x}, y=\bar{y}}. \end{aligned}$$

The Jacobian has different elements depending on whether the system is upload or download constrained at the equilibrium. In the download-constrained case, $\mathbf{J}(\bar{x}, \bar{y})$ is given by

$$\begin{pmatrix} -c & 0 \\ \zeta c - \dfrac{\bar{y}}{\lambda(z - \bar{x}/\lambda)^2} & -\dfrac{1}{z - \bar{x}/\lambda} \end{pmatrix},$$

while in the upload-constrained case, it reads as

$$\begin{pmatrix} -\eta\mu & -\mu \\ \zeta\eta\mu - \dfrac{\bar{y}}{\lambda(z - \bar{x}/\lambda)^2} & \zeta\mu - \dfrac{1}{z - \bar{x}/\lambda} \end{pmatrix}.$$

Let us denote the roots of the characteristic equation (14) by z_1 and z_2. A necessary condition for the local stability is that the real parts of the roots must be negative, i.e.,

$$\mathfrak{Re}(z_i) < 0, \quad i = 1, 2. \tag{15}$$

If the roots satisfy (15) and there is no imaginary part, the equilibrium is approached without oscillations at an exponential rate. If there is an imaginary part, the system is locally asymptotically stable, i.e., the equilibrium is approached with exponentially decaying oscillations around the equilibrium. If $\mathfrak{Re}(z_1) \geq 0$ or $\mathfrak{Re}(z_2) \geq 0$ then the system is both locally and globally unstable.

Application to the P2P system: Here we refer to Figure 2 and apply the local stability results to the steady-state solutions in the three different areas in Figure 2. When the system is download-constrained at the equilibrium, it can be easily verified, recall assumption (4), that the eigenvalues are

$$z_1 = -c < 0, \quad z_2 = \frac{1}{1 - cz} < 0.$$

Thus the system is always locally stable, when it is download-constrained. In Figure 2, the system is locally stable in the area DL. Also in the DL/UL area, where there are two positive equilibria, the DL solution is locally stable.

On the other hand, in the DL/UL area, the upload-constrained solution has eigenvalues of the form,

$$z_1 = \frac{2\lambda}{-\frac{k+z\eta\lambda}{\eta-\zeta} + \sqrt{\frac{\frac{(k\mu+\lambda(\mu z\eta-2))^2(k\mu+\lambda(\mu z\eta-1))^2}{u^2(\eta-\zeta)^2}}{k\mu+\lambda(\mu z\eta-1)}}}$$

and

$$z_2 = \frac{-2\lambda}{\frac{k+z\eta\lambda}{\eta-\zeta} + \sqrt{\frac{\frac{(k\mu+\lambda(\mu z\eta-2))^2(k\mu+\lambda(\mu z\eta-1))^2}{u^2(\eta-\zeta)^2}}{k\mu+\lambda(\mu z\eta-1)}}}.$$

Now, in the UL/DL area we have that $k\mu + \lambda(\mu z\eta - 1) < 0$ (since $\eta < \frac{1}{z}(1/\mu - k/\lambda)$) and $\eta - \zeta < 0$. Hence, the denominator of z_2 is negative and $z_2 > 0$. The upload-constrained equilibrium is thus unstable in the UL/DL area.

In the UL area, where upload-constrained solution is the only equilibrium, it holds that $k\mu + \lambda(\mu z\eta - 1) > 0$ and $\eta - \zeta > 0$. From these it readily follows that $z_2 < 0$. For z_1 we obtain (i) if $\frac{1}{z}\left(\frac{2}{\mu} - \frac{k}{\lambda}\right) < \eta < 1$ then

$$z_1 = -\mu(\eta - \zeta) < 0,$$

and (ii) if $\frac{1}{z}\left(\frac{1}{\mu} - \frac{k}{\lambda}\right) < \eta < \frac{1}{z}\left(\frac{2}{\mu} - \frac{k}{\lambda}\right)$ then

$$z_1 = \frac{-\mu\lambda(\eta - \zeta)}{k\mu + \lambda(z\eta\mu - 1)} < 0.$$

Thus, when the system is in area UL, then the unique equilibrium is locally stable.

In summary, the system has a unique equilibrium (either DL or UL) which is locally stable if $\eta > \frac{1}{z}\left(\frac{1}{\mu} - \frac{k}{\lambda}\right)$. In the area UL/DL, there are two equilibria of which the DL solution is locally stable and the UL solution is unstable.

E. Stochastic model

In addition to a fluid model, we develop a more detailed stochastic model to test the accuracy of the approximative assumptions made for the fluid model. The stochastic model does not utilize the quasi-stationarity approximation to estimate the departure rate of altruistic seeds. Instead, the sojourn time is derived from the stochastically evolving service rate of the system. However, the sharing of the capacity is still idealized, i.e., upload capacity is shared uniformly among the leechers.

Let the capital letters $X(t)$ and $Y(t)$ refer to the random variables corresponding to the number of leechers and seeds, respectively, at time t. We assume that the new peers arrive according to a Poisson process with rate λ. Let $A(t)$, $C(t)$, and $\tilde{D}(t)$ denote the total number of peers arrived, video file transfer completions, and seed departures, respectively, until time t. In addition, let $\tilde{A}(t)$ and $\tilde{C}(t)$ denote the total number of altruistic peers (which continue to upload to leechers after their own transfer phase) that have arrived and that have completed the video file transfer, respectively, until time t. We have immediately the following relations:

$$X(t) = A(t) - C(t),$$
$$Y(t) = \tilde{C}(t) - \tilde{D}(t).$$

Note that $\tilde{A}(t)$ is derived from $A(t)$ by including each arriving peer in $\tilde{A}(t)$ with probability ζ according to an independent Boolean trial. Below we show that in our model also $C(t)$, $\tilde{C}(t)$, and $\tilde{D}(t)$ are determined from the arrival process $A(t)$ (and its derivative $\tilde{A}(t)$).

While the playback rate w is deterministic and constant, the transfer rate $R(t)$ is random and time-varying depending on the dynamics of the whole file sharing system. More precisely, $R(t)$ depends on the number of peers, the access link rates, and the efficiency of the video file sharing. The total upload capacity at time t is equal to $u(\eta X(t) + Y(t) + k)$. We assume that this upload capacity is divided evenly among the leechers and there are no bottlenecks in the core network. Thus, we get the following expression for the transfer rate (whenever there are leechers, i.e., $X(t) > 0$):

$$R(t) = \min\{d, u(\eta + \frac{Y(t) + k}{X(t)})\}. \tag{16}$$

The transfer phase of a peer lasts until the entire video file has been retrieved. Since the transfer rate $R(t)$ in our model is the same for all peers, a peer that arrived at time T_a completes the transfer phase at time T_c, where

$$T_c = \inf\{t \geq T_a \mid \int_{T_a}^{t} R(s)\,ds \geq m\}. \tag{17}$$

Vice versa, a peer that completes the transfer phase at time T_c arrived at time T_a, where

$$T_a = \sup\{t \leq T_c \mid \int_{t}^{T_c} R(s)\,ds \geq m\}. \tag{18}$$

It follows from (18) that

$$C(t) = A(\sup\{t' \le t \mid \int_{t'}^{t} R(s)\,ds \ge m\}),$$

$$\tilde{C}(t) = \tilde{A}(\sup\{t' \le t \mid \int_{t'}^{t} R(s)\,ds \ge m\}).$$

In our model there is no guarantee that the length of the transfer phase, $T_c - T_a$ (derived from (17)), is less than the length z of the playback phase with the full playback rate. If this is not the case, we assume that the playback is delayed so that the peer stays in the system until the end of the transfer phase and departs from the system just immediately thereafter. Thus, we have

$$\tilde{D}(t) = \min\{\tilde{A}(t-z), \tilde{C}(t)\}, \tag{19}$$

which completes our model.

IV. Steady state and scalability

By experimenting with the two models, we obtained that the fluid model closely approximates the mean values of the stochastic model when η is high enough but shows a qualitatively different behavior when η is below a certain threshold. In this section we show how insights from the numerical simulation of the stochastic model and the analytical results can be combined to provide a meaningful synthesis of the system performance in the whole (η, ζ) parameter region as well as an interpretation of the results from the point of view of system scalability.

A. Steady-state synthesis

According to extensive numerical experiments (to be summarized in Section V), we have the following conclusions regarding the steady-state behavior of the stochastic model:

(i) If

$$\eta > \frac{1}{z}\left(\frac{1}{\mu} - \frac{k}{\lambda}\right), \tag{20}$$

the equilibrium transfer rate for a leecher exceeds the playback rate w (so that the the playback quality is sufficient for the VoD application). If additionally (8) is satisfied, then the system is download-constrained, and the numbers of leechers and seeds are well estimated by $\bar{x}_d$ and $\bar{y}_d$ given in (7). Otherwise, the system is upload-constrained, and the numbers of leechers and seeds are well estimated by $\bar{x}_u$ and $\bar{y}_u$ given in (9).

(ii) If

$$\eta \le \frac{1}{z}\left(\frac{1}{\mu} - \frac{k}{\lambda}\right), \tag{21}$$

then the equilibrium transfer rate for a leecher stays below the playback rate w resulting in playback quality problems. In this case, in the stochastic model peers have to stay longer in the system than the actual viewing time and (19) gives us that $Y(t) = 0$. In our fluid model this situation can be enforced by evaluating the steady-state solution with the condition $\zeta = 0$. Then, the system is always upload-constrained (see Figure 2) and

the numbers of leechers and seeds are well estimated by formula (9) applied with $\zeta = 0$ leading to estimates

$$\begin{cases} \bar{x}_0 &= x_u|_{\zeta=0} &= \frac{\lambda}{\eta}\left(\frac{1}{\mu} - \frac{k}{\lambda}\right), \\ \bar{y}_0 &= y_u|_{\zeta=0} &= 0. \end{cases} \tag{22}$$
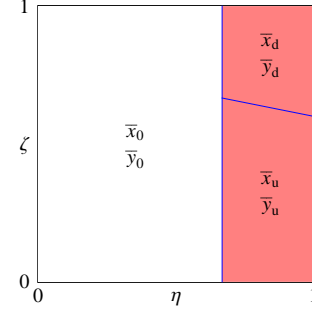
These conclusions are summarized in Figure 3.

Fig. 3. Steady-state synthesis.

Note that the altruism parameter ζ has no effect on the threshold (20). However, the per-leecher service rate increases with ζ (up to the download rate limitation d) implying that the system offers better viewing quality. This can be verified by noting that $\bar{x}$ decreases and $\bar{y}$ increases in (9) as ζ increases, also see Figure 6, and that the per-leecher service rate is given by (16) where $X(t) = \bar{x}$ and $Y(t) = \bar{y}$. In addition, we see from (20) that $\frac{1}{\mu} - \frac{k}{\lambda} < z$ is a necessary condition and $\frac{1}{\mu} - \frac{k}{\lambda} < 0$ is a sufficient condition for good playback quality.

B. Scalability

As explained in the previous section IV-A, the system seems to have a sufficient playback quality if and only if condition (20) is satisfied. We say that the system is *scalable* if (20) is satisfied for all λ. Thus, scalability is equivalent to the requirement that

$$\eta > \frac{1}{z\mu}. \tag{23}$$

A necessary condition for scalability is clearly that

$$z > 1/\mu, \tag{24}$$

which is equivalent to $u > w$. Note also that neither the altruism parameter ζ nor the number k of permanent seeds affect the scalability property.

On the other hand, we see from (20) that if the system is *not* scalable (so that $1/\mu \ge \eta z$), good playback quality is achieved up to

$$\lambda < \frac{k}{\frac{1}{\mu} - \eta z}. \tag{25}$$

C. Effect of permanent seeds and altruism

A service provider may have the possibility of maintaining a number of permanent seeds, k, in the system. In general, it would seem intuitive that increasing k would somehow help the system to perform better. However, as can be seen from

(23), to achieve scalability does not depend on k. But in the non-scalable case, by (25), the limit for λ up to which playback quality is sufficient increases linearly with k.

While we have showed that neither the number of permanent seeds k nor the level of altruism ζ had any effect on the scalability of the system, it does not follow that these are unnecessary parameters. In fact, ζ is a determining factor in whether the system reaches a download or upload constrained steady state. Given an efficiency η satisfying (20), one can always find a needed level of altruism $\zeta < 1$ so that the steady state is download-constrained. In a download-constrained steady state, each seed in the system does not need to use all of its upload capacity, and in fact the service provider could drop some or all of its permanent seeds from the system without repercussions. In the case of a YouTube-like VoD service provider, the required upload bandwidth per video is a significant portion of costs [14]. Thus even a slight reduction in needed upload bandwidth can be significant.

V. NUMERICAL RESULTS AND VALIDATION

In this section, we validate the accuracy of the fluid model against simulations from the stochastic model of Section III-E and traces from a more realistic BitTorrent simulator implementing a windowing algorithm. We first study the system dynamics and then consider the steady-state performance. Similarly as in the steady-state analysis, we focus on the behavior of the system in the (η, ζ) parameter space, cf., Figures 2 and 3.

A. Youtube scenario

In the following tests we consider a viewing scenario where the parameters correspond to a typical YouTube setting. The users are viewing a video file consisting of 800 pieces each 32kB in size. The video coding rate is $w = 300$ kbit/s, and thus the viewing time is $z = 682$ s. The upload and download bandwidths of the users are $u = 512$ kbits/s and $d = 1024$ kbit/s (typical asymmetric ADSL subscriber rates). Also, we assume that $k = 1$ (one permanent seed).

In the simulations, new users arrive according to a Poisson process with rate $\lambda = 0.2$ peers/s. The simulation of the stochastic model has been implemented in discrete time with a time granularity of 1 s. Thus, the arrivals in a time slot obey a Poisson distribution with mean λ. In both the fluid and the stochastic model, the initial buffering time was assumed negligible. As an example of a realistic P2P VoD system we used traces from a simulation of BitTorrent with the Stretching Window algorithm, as described in [10]. In the simulations the initial buffer was set equal to 4 pieces (each 32kB), which corresponded to an initial delay typically ranging between 3 to 10 seconds. This was empirically determined to provide a sufficient playback quality for the peers.

B. Validation of system dynamics

Here we give results to validate the accuracy of the fluid model dynamics (6) and the use of the quasi-stationarity approximation in its derivation. The idea is to compare the
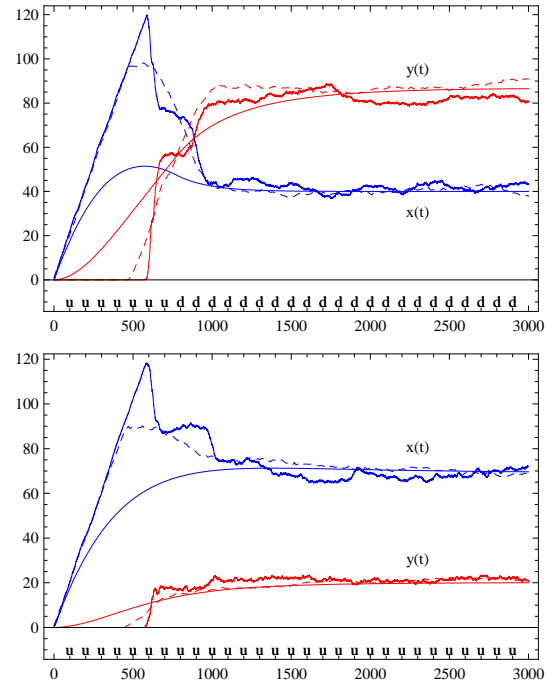

Fig. 4. Comparison of the fluid model (solid smooth line) against the stochastic model (dashed line) and the BitTorrent simulation (solid jagged line) with $\zeta = 0.9$ (upper panel) and $\zeta = 0.3$ (lower panel).

dynamic behavior of (6) with the results both from the stochastic model and the BitTorrent simulations.

We first study two examples where the users are first altruistic with $\zeta = 0.9$ and then non-altruistic with $\zeta = 0.3$. The parameters correspond to the situations where condition (20) holds, i.e., the system has a stable and unique equilibrium. The results are shown in Figure 4, where the upper panel corresponds to the case $\zeta = 0.9$ and the lower panel to the case $\zeta = 0.3$. The figure shows the time evolution of the fluid model (solid smooth line), stochastic model (dashed line) and the BitTorrent simulation (solid jagged line) for both the number leechers $x(t)$ and the number of seeds $y(t)$ when the system is initially empty, i.e., $x(0) = 0$ but $y(0) = 0$. The simulations represent averages over 20 sample paths. Finally, the letters "u" and "d" indicate whether the service rate of the system at each time instant is upload or download limited in the fluid model. In the figure, we have numerically determined the value of the efficiency parameter η so that it matches best the equilibrium values of the actual BitTorrent simulations. This resulted in $\eta = 0.85$ for $\zeta = \{0.9, 0.3\}$.

When $\zeta = 0.9$ (see upper panel of Figure 4), one can observe that the system is initially upload-constrained. However, after reaching equilibrium it is download-constrained, i.e., the total number of seeds and leechers in the system yields so much upload capacity that each leecher is constrained only by the amount of download bandwidth available. As ζ is decreased the number of seeds begins to decrease, which decreases the service capacity available for leechers, and with $\zeta = 0.3$ (see lower panel of Figure 4) the available capacity
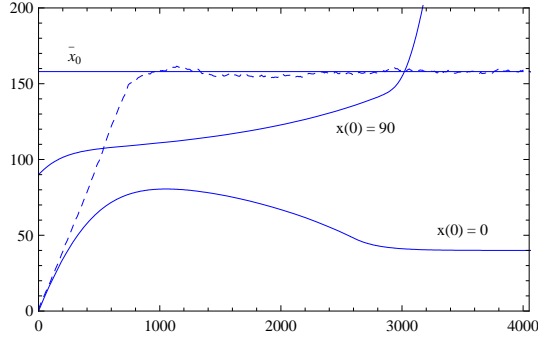
Fig. 5. Comparison of the fluid model (solid smooth lines) against the stochastic model (dashed line) with $\eta = 0.5$ and $\zeta = 0.8$.

is now limited by the uploading bandwidth of the system.

With regard to the accuracy of the fluid model we can observe that the equilibrium values of the system are very accurately the same between the stochastic model and the fluid model, and also the BitTorrent simulation. However, the dynamics of the fluid model (solid smooth lines) are somewhat smoother than of the corresponding discrete time simulation of the stochastic model (dashed lines). This can be partly explained by the fact that the model assumes that seeds are continuously generated (recall the quasi-stationary approximation in the use of Little's result), while in the simulated system seeds start accumulating only after the leechers have received enough service to complete the download. Comparing the BitTorrent simulation and the stochastic model we can observe a slight difference in the dynamics (see upper panel in Figure 4): the number of seeds increases in two phases in the BitTorrent simulation (solid jagged lines) while the increase is smoother for the stochastic model. This may be related to the fact that in the stochastic model we assume that all leechers are uniformly able to provide pieces of the file to other leechers (idealized service assumption), while in the actual BitTorrent simulation this is not always true due to the imperfection of the actual policies used to select peers/pieces.

Finally, we illustrate the dynamics when η and ζ are such that condition (20) does not hold and we are in the DL/UL area of Figure 2. Figure 5 shows the dynamics of the number of leechers for the case when $\eta = 0.5$ and $\zeta = 0.8$ (recall that in the stochastic model the number of seeds goes to zero in this parameter area). In the figure, the horizontal line $\bar{x}_0$ corresponds to the estimated steady state value from (22), the dashed line corresponds to the stochastic model and the two solid lines corresponds to two solutions from the fluid model with two different initial conditions, $x(0) = 0$ and $x(0) = 90$, respectively. The two fluid model solutions with different initial conditions indicate the dynamic behavior of the system that depending on the initial condition the system may arrive at the stable DL solution (when $x(0) = 0$) or may approach the unstable UL solution (when $x(0) = 90$). On the other hand, in the stochastic model simulations the number of leechers stabilizes at a value well approximated by $\bar{x}_0$.
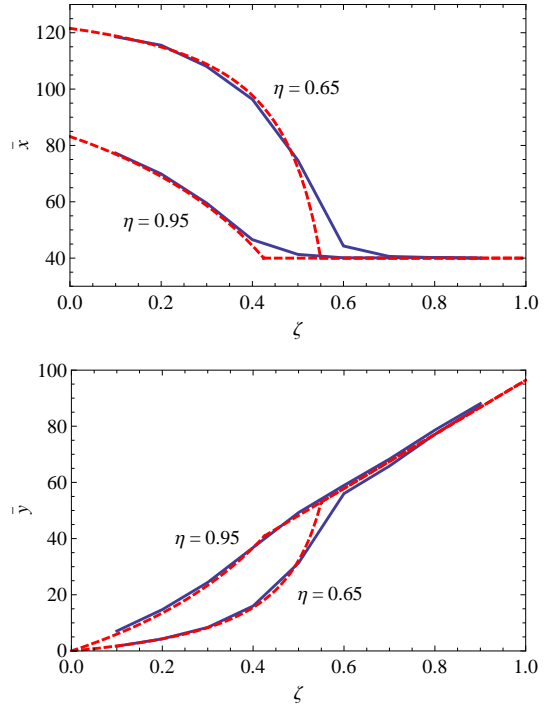

Fig. 6. Comparison of the steady-state estimates of the fluid model (solid red line) against the stochastic model (solid blue line) for number of leechers (upper panel) and seeds (lower panel).

C. Validation of steady-state estimates and scalability

We first analyze the steady-state performance in view of the steady-state synthesis results in Section IV-A. The accuracy of the steady-state estimates given by the fluid model are compared to the steady-state mean values of the stochastic model. The results are given in Figure 6, which depicts the number of leechers (upper panel) and the number seeds (lower panel) as a function of ζ for $\eta = 0.65$ and $\eta = 0.95$. In the figures, the solid lines correspond to simulated estimates from the stochastic model and the dashed lines correspond to the analytical estimates of the fluid model. The simulated results for the stochastic model were obtained (for each value of ζ) by simulating the stochastic model for 15 000 s and using the samples from the last 10 000 s to compute an estimate of the mean number of leechers and seeds. The final result is an average of 10 such realizations. As can be seen, the estimates of the fluid model very closely match the results from the simulations.

Next we focus on validating the scalability results in Section IV-B and study the ratio of the number of leechers to the number of seeds when λ is increased. The results are shown in Figure 7, where the solid lines correspond to simulated estimates from the stochastic model and the dashed lines correspond to the analytical estimates of the fluid model. Note that the leecher-to-seed ratio is shown on logarithmic scale to highlight the differences between the cases. By the condition (23), the system is scalable if $\eta > 0.59$ with the present parameter values, i.e., the performance does not depend on
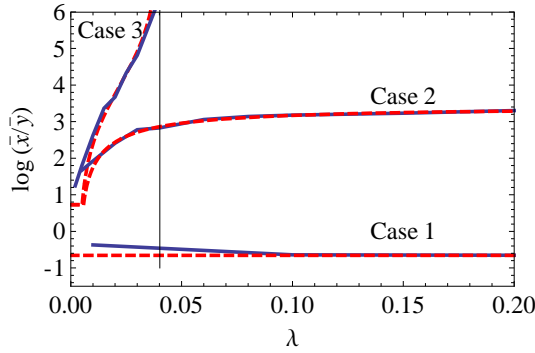
Fig. 7. Validation of the scalability as a function of λ. Cases 1 and 2 are scalable, while Case 3 is not.

λ. This is illustrated in the figure by Cases 1 and 2, where we have used $\eta = 0.65$. In Case 1, the altruism parameter $\zeta = 0.8$ and the system stays download-constrained independent of λ and by (7) the leecher-to-seed ratio is also constant. In Case 2, the altruism parameter $\zeta = 0.2$ and the steady state evolves so that for small λ the system is download-constrained but then it switches to being upload-constrained. However, the leecher-to-seed ratio tends to a constant, as predicted by (9). On the other hand, in Case 3 we have $\eta = 0.55$ and by (21) the system is stable only up to $\lambda = 0.043$ (illustrated by the vertical line in the figure). As can be seen, the above behavior, as predicted by the analytical results of the fluid model, matches very well the simulated results from the stochastic model. The numerical accuracy of the leecher-to-seed ratio is also well approximated by the results from the fluid model, especially for higher values of λ. For very small values of λ, also $\bar{x}$ and $\bar{y}$ remain small and the approximations of the fluid model result in some inaccuracies compared with the stochastic model.

VI. CONCLUSIONS

We have derived a fluid model for P2P on-demand streaming of stored content, where the coupled transfer and playback phases are explicitly taken into account, unlike in [2]. Our model also makes what we feel is a more realistic assumption about peer altruism, where altruistic peers stay on as seeds only as long as they are watching the video. The model consists of a system of differential equations for characterizing the time-evolution of the mean number of leechers and seeds in the system. Our analytical results include the analysis of the uniqueness and the local stability of the steady-state solution.

Based on the analytical results and thorough numerical validations, we have concluded the following. The efficiency parameter η plays a central role for the scalability of the system. By scalability we mean that the streaming rate of the system is sufficient to guarantee good viewing quality independent of the arrival rate λ. In analysis of file sharing systems, it is often argued that $\eta \approx 1$ in practice. However, in streaming applications this may not be the case due its strong dependence on the window control algorithm. We showed that η must be greater than a certain threshold in order to achieve scalability. Also, a necessary condition for scalability is that

both the download rate d and the upload rate u must be greater than the viewing rate w. If the system is not scalable, then acceptable viewing quality is achieved only up to a given λ. This bound can be made higher by, e.g., increasing the number of permanent seeds k.

We also showed that if the system is scalable, a sufficient level of playback phase altruism ζ implies that the system reaches a download-constrained steady state rather than an upload-constrained one, which we believe would result in reducing the bandwidth needed by a VoD service provider, as well as the associated costs.

The potential for future research in this area is rich, including relaxation of some of the modeling assumptions (non-homogeneous upload/dowload rates, finite playback buffer, dynamic model for η, allowing for peers to abort their downloads). The actual P2P systems useful for VoD, such as windowing BitTorrent, remain open topics, especially making them as ISP-friendly as possible. Finally, analyzing the stochastic and fluid models and applying the results to create a load balancing or system dimensioning tool for a P2P VoD service provider is also an intersting topic for future research.

REFERENCES

[1] B. Cohen, "Incentives build robustness in bittorrent," in *Proc. of the 1st Workshop on Economics of Peer-to-Peer Systems*, 2003.
[2] K. Parvez, C. Williamson, A. Mahanti, and N. Carlsson, "Analysis of BitTorrent-like protocols for on-demand stored media streaming," in *Proc. of ACM SIGMETRICS*, 2008, pp. 301–312.
[3] T. Bonald, L. Massoulié, F. Mathieu, D. Perino, and A. Twigg, "Epidemic live streaming: Optimal performance trade-offs," in *Proc. of ACM SIGMETRICS*, 2008, pp. 325–336.
[4] R. Kumar, Y. Liu, and K. Ross, "Stochastic fluid theory for p2p streaming systems," in *Proc. of IEEE INFOCOM*, 2007, pp. 919–927.
[5] D. Qiu and R. Srikant, "Modeling and performance analysis of BitTorrent-like peer-to-peer networks," in *Proc. of ACM SIGCOMM*, 2004, pp. 367–378.
[6] X. Yang and G. de Veciana, "Service capacity of peer to peer networks," in *Proc. of IEEE INFOCOM*, 2004, pp. 2242–2252.
[7] D. Qiu and W. Sang, "Global stability of peer-to-peer file sharing systems," *Computer Communications*, vol. 31, no. 2, pp. 212–219, 2008.
[8] F. Lo Piccolo, G. Neglia, and G. Bianchi, "The effect of heterogeneous link capacities in BitTorrent-like file sharing systems," in *Proc. of HOT-P2P*, 2004, pp. 40–47.
[9] Y. Yue, C. Lin, and Z. Tan, "Analyzing the performance and fairness of BitTorrent-like networks using a general fluid model," *Computer Communications*, vol. 29, no. 18, pp. 3946–3956, 2006.
[10] P. Savolainen, N. Raatikainen, and S. Tarkoma, "Windowing BitTorrent for video-on-demand: Not all is lost with tit-for-tat," in *Proc. of IEEE GLOBECOM*, 2008.
[11] S. Tewari and L. Kleinrock, "Analytical model for BitTorrent-based live video streaming," in *Proc. of IEEE CCNC*, 2007, pp. 976–980.
[12] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *Proc. of USENIX IMC*, 2007, pp. 1–14.
[13] D. Liberzon and A. Morse, "Basic problems in stability and design of switched systems," *IEEE Control Systems Magazine*, vol. 19, no. 5, pp. 59–70, Oct 1999.
[14] C. Huang, J. Li, and K. Ross, "Can internet video-on-demand be profitable?" in *Proc. of ACM SIGCOMM*, 2007, pp. 133–144.