On the Impact of Bounded Rationality in Peer-to-Peer Networks

Hyunggon Park, *Member, IEEE*, and Mihaela van der Schaar, *Senior Member, IEEE*

***Abstract*—In this letter, we consider peer-to-peer (P2P) networks, where multiple peers are interested in sharing their content. In the considered P2P system, autonomous and self-interested peers use a Markov Decision Process (MDP) framework to determine their upload bandwidth allocations, which maximize their individual utilities. This framework enables the peers to make foresighted decisions on their bandwidth allocations, by considering the future impact of their decisions. In this letter, we focus on the impact of the peers' bounded rationality on their resource reciprocation strategies and ultimately, on their achievable utilities. Specifically, we consider peers who have only a limited ability to model the other peers' strategies for resource reciprocation, and study how this impacts their own decisions.**

***Index Terms*—Bounded rationality, peer-to-peer (P2P) networks, resource reciprocation.**

I. Introduction

We consider the resource reciprocation among peers in data-driven P2P systems such as CoolStreaming [1] or BitTorrent [2], which adopt pull-based techniques [1]. While this approach has been successfully deployed in real-time multimedia streaming or file-sharing applications over P2P networks, little is known about determining optimal resource reciprocation strategies among self-interested peers, which aim to maximize their own utilities (e.g., download rates) based on their heterogeneous capabilities, in repeatedly interacting environment.

Several resource reciprocation strategies for self-interested peers in P2P networks have been proposed in e.g., [2]–[5]. To provide incentives for the peers that contribute their resources (e.g., bandwidth, content, etc.), exchange-based incentive mechanisms including tit-for-tat (TFT) strategies in BitTorrent systems are proposed [2]–[4]. While these approaches provide appropriate incentives for peers' contributions, they are *myopic*, as they do not explicitly consider repeated interactions among the peers. In order to take into account the repeated interactions, an evolutionary instantiation of the Generalized Prisoner's Dilemma is proposed [5]. However, this work studies only a simple P2P interaction scenario, where peers have a limited set of actions—either allowing download or ignoring download requests. Hence, this work does not provide solutions for how to optimally divide each peer's available resources (i.e., determine their level of cooperation). Unlike these approaches, in [6], *foresighted* policies based on an MDP framework are proposed, which enable the peers to determine their resource allocations in a way that maximizes their long-term utilities. Thus, the foresighted policies can improve the performance of the participating peers, compared to existing solutions such as TFT [2] or BitTyrant [7]. However, this work does not consider the complexity and memory requirements to find the foresighted policies, which may become critical for practical implementation of the MDP-based solutions.

Since peers generally have in practice memory and/or computational constraints, they have limited knowledge of the other players' behavior and limited ability to analyze their environment, i.e., peers are *boundedly rational* [8]. Hence, in this letter, we investigate how the computation and memory limitations of peers (i.e., their bounded rationality) affect their foresighted policies and the resulting performance. Specifically, we quantify the impact of the bounded rationality on the accuracy of long-term utilities. This analysis enables each peer to minimize the complexity and/or memory requirements, while achieving the desirable accuracy of the utilities.

II. Resource Reciprocation Based on MDP

A. Resource Reciprocation in Groups

In P2P networks, peers share content with each other in their *groups* [1], while negotiating the amount of resources which they will provide to each other. We denote a group of N_{C_i} peers associated with peer i by C_i, i.e., $C_i = \{1, \ldots, N_{C_i}\}$. Note that peer $k \in C_i$ also has its own group C_k that includes peer i. Thus, peers can belong to multiple groups. For more information about the problem set-up, the interested reader is referred to [6].

B. MDP Based Resource Reciprocation

A peer i determines its resource allocations (i.e., actions) to the associated peers in C_i. The actions of a peer i are determined using a foresighted policy obtained based on an MDP [6], which enables the peer to maximize its cumulative discounted expected rewards (CDERs). Specifically, for a peer i in its group C_i, an MDP is a tuple $\langle \mathbf{S}_i, \mathbf{A}_i, P_i, R_i \rangle$, where $\mathbf{S}_i, \mathbf{A}_i, P_i$, and R_i denote the state space, the action space, a state transition probability (STP) function, and a reward function.

1) State Space $\mathbf{S}_i$*:* A state of peer i represents the set of received resources from the peers in C_i. Let $(x_{1i}, \ldots, x_{N_{C_i}i})$ be a set of resources that peer i receives from its associated peers in C_i, where $x_{ki} (0 \le x_{ki} \le L_k)$ denotes the resources provided by peer $k \in C_i$ with its maximum available upload bandwidth L_k to peer i [2]. A function $\psi_{ik}^{(n_{ik})}$ of peer i maps x_{ki} into one

[1]For example, a group can be a partnership in [1], or a swarm in [2].

[2]Equivalently, x_{ki} represents the download rates of peer i from peer k.

Manuscript received January 30, 2009; revised April 08, 2009. First published May 02, 2009; current version published June 05, 2009. This work was supported by the NSF Career and NSF CCF 0830556. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. James E. Fowler.

H. Park was with the Electrical Engineering Department, University of California, Los Angeles, CA 90024 USA. He is now with the Signal Processing Laboratory (LTS4), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland (e-mail: hyunggon.park@epfl.ch).

M. van der Schaar is with the Electrical Engineering Department, University of California, Los Angeles, CA 90024 USA (e-mail: mihaela@ee.ucla.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2009.2022146

of n_{ik} discrete values, referred to as *state descriptions*, and it is defined as $\psi_{ik}^{(n_{ik})}(x_{ki}) = s_{ik}^l$ for $x_{ki} \in \Delta_{ik}^l$ $(1 \le l \le n_{ik})$. Δ_{ik}^l is the lth segment among n_{ik} segments, and it is denoted by the lth state description s_{ik}^l of peer i for peer k. We assume that the available bandwidth is uniformly divided, i.e., Δ_{ik}^l represents $[(l-1)\cdot L_k/n_{ik}\ ,\, l \cdot L_k/n_{ik})$. Hence, the state space of peer i can be expressed as

$$\mathbf{S}_i = \left\{ s_i = (s_{i1}, \ldots, s_{iN_{C_i}}) \middle| s_{ik} = \psi_{ik}^{(n_{ik})}(x_{ki}), \text{for all} k \in C_i \right\}.$$

In this letter, the bounded rationality of peer i is represented by the numbers of state descriptions n_{ik} for its associated peer $k \in C_i$, which will determine its ability to model the evolution of its resource reciprocation with peer k. We study its impact on performance in Section III.

2) Action Space $\mathbf{A}_i$: An action of peer i is its resource allocation to the peers in C_i. Hence, the action space $\mathbf{A}_i = A_i \times \cdots \times A_i$ of peer i in C_i can be expressed as

$$\mathbf{A}_i = \left\{ \mathbf{a}_i = (a_{i1}, \ldots, a_{iN_{C_i}}) \middle| \, 0 \le a_{ik} \le L_i, k \in C_i \right\},$$

where $a_{ik} \in A_i$ denotes the resources allocated by peer i to peer $k \in C_i$ and $\sum_{k \in C_i} a_{ik} \le L_i$. Hence, peer i's action a_{ik} to peer k becomes peer k's received resources from peer i, i.e., $a_{ik} = x_{ik}$. We assume that the available resources (i.e., upload bandwidth) of peers are decomposed into "units" of bandwidth.

3) State Transition Probability (STP) $P_{\mathbf{a}_i}(s_i, s_i')$: An STP represents the probability that by taking an action, a peer will transit into a new state, i.e., $\mathbf{S}_i \times \mathbf{A}_i \times \mathbf{S}_i \to [0,1]$. Hence, given a state $s_i = (s_{i1}, \ldots, s_{iN_{C_i}}) \in \mathbf{S}_i$ at time t, an action $\mathbf{a}_i \in \mathbf{A}_i$ of peer i can lead to another state $s_i' \in \mathbf{S}_i$ at $t+1$ with probability $P_{\mathbf{a}_i}(s_i, s_i') = \Pr(s_i' | s_i, \mathbf{a}_i)$. The STPs can be efficiently estimated based on past resource reciprocations [6].

4) Reward R_i: A reward of a peer i represents the total download rates from C_i. Since the received resources of peer i are represented by its state $s_i \in \mathbf{S}_i$, the reward from s_i can be expressed as

$$R_i(s_i) = R_i(s_{i1}, \ldots, s_{iN_{C_i}}) = \sum_{k \in C_i} r(s_{ik})$$

where $r(s_{ik})$ is a random variable representing the download rates in $s_{ik} = \psi_{ik}^{(n_{ik})}(x_{ki})$.

C. Foresighted Resource Reciprocation Policy

A foresighted peer i's CDERs are expressed as

$$R_i^{CDER}(s_i^{(t)}) \triangleq \sum_{t=t_c+1}^{\infty} \gamma_i^{(t-(t_c+1))} \cdot E\left[R_i(s_i^{(t)})\right] \quad (1)$$

where peer i is in state $s_i^{(t)} = (s_{i1}, \ldots, s_{iN_{C_i}})$ at time $t = t_c$ given a discount factor γ_i. Thus, the peer i takes its action $\mathbf{a}_i^*$ that maximizes $R_i^{CDER}(s_i^{(t)})$ in (1), i.e.,

$$\mathbf{a}_i^* = \arg\max_{\mathbf{a}_i \in \mathbf{A}_i} R_i^{CDER}(s_i^{(t)}) \quad \text{subject to} \quad \sum_{k \in C_i} a_{ik} \le L_i. \quad (2)$$

Solutions to the optimization problem in (2) determine an optimal policy π_i^*, which maps each state $s_i \in \mathbf{S}_i$ to a corresponding optimal action $\mathbf{a}_i^*$, i.e., $\pi_i^*(s_i) = \mathbf{a}_i$ for all $s_i \in \mathbf{S}_i$, where action a_{ik} to peer k is denoted by $a_{ik} = [\pi_i^*(s_i)]_k$. π_i^* can be obtained using well-known methods such as value/policy iteration [9]. Note that the discount factor γ_i in (1) can alternatively represent the belief of the peer i about the validity of the expected future rewards [6]. Thus, we assume that the discount factor of each peer is determined using its past experiences, reputation of their associated peers [6], or a tolerable accuracy of CDERs, as will be discussed later.

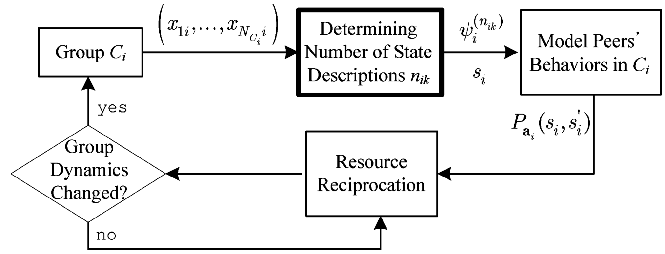

Fig. 1. Resource reciprocation of peer i based on the MDP.

The resource reciprocation process of peer i based on the MDP is depicted in Fig. 1. In this letter, we focus on how peer i can determine the number of state descriptions n_{ik} for peer k, while explicitly considering the impact of n_{ik} on the accuracy of the achieved CDERs. This is highlighted in Fig. 1.

III. Bounded Rationality of Peers

A. Bounded Rationality: Finite Number of State Descriptions

As discussed in Section II-B, a peer's received resources from its associated peers are captured by its state, and the boundedly rational peers can use only a finite number of state descriptions for their states. Moreover, a reward from s_{ik} is represented by a random variable $r(s_{ik})$. Since no prior information about the action of peer k is available, we assume that $r(s_{ik}^l)$ for $1 \le l \le n_{ik}$ is a uniform random variable, where its mean and variance are given by $E(r(s_{ik}^l)) = (2l-1)L_k/2n_{ik}$ and $\sigma^2(r(s_{ik}^l)) = 1/12\,(L_k/n_{ik})^2$, respectively.

Thus, it is obvious that using more state descriptions (i.e., finer states) enables each peer to compute the actual CDERs more accurately. However, increasing the number of state descriptions also leads to higher computational complexity to find the optimal policy [9]. Therefore, it is important for each peer to minimize the number of state descriptions, while achieving a tolerable accuracy of the actual CDERs. To quantify how accurately the CDERs computed based on a finite number of state descriptions represent the actual CDERs, we use the variance of the computed CDERs. We focus on quantifying the impact of the accuracy improvement on the CDERs by using more state descriptions.

B. Impact of State Granularity on CDER Accuracy

In this section, we analytically quantify the impact of the state granularity on CDERs of a peer i for peer k based on the MDP described in Section II [3].

Let π_i^* be an optimal policy of peer i with n state descriptions for peer k. The corresponding STPs are expressed as an $n \times n$ matrix $\mathbf{P}_{(n)} = [\mathbf{p}_1, \ldots, \mathbf{p}_n]^T$, where $\mathbf{p}_l^T = [p_{l1}, \ldots, p_{ln}]$ for $p_{l_1 l_2} = P_{a_{ik}^*}(s_{ik}^{l_1}, s_{ik}^{l_2})$ $(1 \le l_1, l_2 \le n)$ and a_{ik}^* is determined by π_i^*. As shown in [6], the CDER of peer i from peer k in s_{ik}^l

[3]For notation simplicity, we omit the subscript of the number of state descriptions of peer i for peer k in the following analysis.

is denoted by $J^*_{(n)}(s^l_{ik})$ for $1 \le l \le n$, and it can be expressed as

$$J^*_{(n)}(s^l_{ik}) = \mathbf{p}_l^T \mathbf{r}_{(n)} + \frac{\gamma_i}{1-\gamma_i}\boldsymbol{\nu}^T \mathbf{r}_{(n)} \tag{3}$$

where γ_i ($0 \le \gamma_i < 1$) denotes a discount factor of peer i, and $\mathbf{r}_{(n)} = \left[r(s^1_{ik}), \dots, r(s^n_{ik})\right]^T$ denotes a reward vector. $\boldsymbol{\nu}^T$ is a steady state probability vector of $\mathbf{P}_{(n)}$, i.e., $\lim_{h\to\infty}[\mathbf{P}_{(n)}]^h = \mathbf{1}\boldsymbol{\nu}^T$ where $\mathbf{1} = (1, \dots, 1)^T$. Hence, the CDER of peer i from peer k can be computed based on its STP and steady state probability vector. We assume that the discount factor γ_i of peer i for MDP is predetermined.

To study the impact of the bounded rationality of peer i on the accuracy of its CDERs, we consider the case where peer i uses m ($m \le n$) state descriptions for peer k. The impact of the bounded rationality on the accuracy of obtained CDERs are stated in Proposition 1.

Proposition 1: If each state for a peer is refined using $(n/m) \in \mathbb{N}$ state descriptions, the accuracy of CDERs is improved by a factor of at least $(n/m)^2$ and at most $(n/m)^3$.

Proof: Let $\mathbf{P}_{(m)}$ be a $m \times m$ state transition probability matrix obtained by merging $n/m \in \mathbb{N}$ state descriptions of $\mathbf{P}_{(n)}$ into one state description, i.e., $s^{l\,\prime}_{ik} = (s^{(n/m)(l-1)+1}_{ik} \cup \cdots \cup s^{(n/m)l}_{ik})$ for $1 \le l \le m$. Hence, given an initial distribution $\boldsymbol{\mu}^T = (\mu_1, \dots, \mu_n)$, where peer k is in s^l_{ik} with probability μ_l, $\mathbf{P}_{(m)}$ can be expressed as

$$\mathbf{P}_{(m)} = [\mathbf{p}'_1, \dots, \mathbf{p}'_m]^T \tag{4}$$

where $\mathbf{p}'^T_l = [p'_{l1}, \dots, p'_{lm}]$ and each element of $\mathbf{P}_{(m)}$ is determined by

$$p'_{l_1 l_2} = \frac{\sum_{l=(n/m)(l_1-1)+1}^{(n/m)l_1} \mu_l \left(\sum_{h=(n/m)(l_2-1)+1}^{(n/m)l_2} p_{lh}\right)}{\sum_{l=(n/m)(l_1-1)+1}^{(n/m)l_1} \mu_l} \tag{5}$$

for $1 \le l_1, l_2 \le m$. The corresponding policy $\pi^{*\prime}_i$ is $\pi^{*\prime}_i(s^{l\,\prime}_{ik}) = \sum_{h=1}^{n/m} \pi^*_i(s^{(n/m)(l-1)+h}_{ik})\mu_{(n/m)(l-1)+h}$ for $1 \le l \le m$. Thus, the resulting CDERs from $s^{l\,\prime}_{ik}$ can be expressed as

$$J^*_{(n)}(s^{l\,\prime}_{ik}) = \frac{\sum_{h=(n/m)(l-1)+1}^{(n/m)l} \mu_h J^*_{(n)}(s^h_{ik})}{\sum_{h=(n/m)(l-1)+1}^{(n/m)l} \mu_h}. \tag{6}$$

Note that peer i with m state descriptions, however, cannot achieve $J^*_{(n)}(s^{l\,\prime}_{ik})$ in (6), as it cannot discriminate between $s^{(n/m)(l-1)+1}_{ik}, \dots, s^{(n/m)}_{ik}$, due to its bounded rationality for describing states. Rather, the peer i considers the reward from $s^{l\,\prime}_{ik}$ as $r(s^{l\,\prime}_{ik})$, where its mean and variance are given by $E(r(s^{l\,\prime}_{ik})) = L_k/2n \cdot n/m(2l-1)$ and $\sigma^2(r(s^{l\,\prime}_{ik})) = 1/12\,(L_k/m)^2$ for $1 \le l \le m$. We denote the steady state probability vector of $\mathbf{P}_{(m)}$ by $\boldsymbol{\nu}'^T$. Therefore, the effect of the bounded rationality on the accuracy of CDERs can be quantified by comparing the variances of $J^*_{(n)}(s^{l\,\prime}_{ik})$ and $J^*_{(m)}(s^{l\,\prime}_{ik})$.

Without loss of generality, we can compare the two quantities $\sigma^2(J^*_{(n)}(s^{1\,\prime}_{ik}))$ and $\sigma^2(J^*_{(m)}(s^{1\,\prime}_{ik}))$. Based on (3) and (6),

$$\begin{aligned}
&\sigma^2(J^*_{(n)}(s^{1\,\prime}_{ik})) \\
&\approx \sigma^2\left(\frac{\sum_{l=1}^{n/m}\mu_l \mathbf{p}_l^T \mathbf{r}_{(n)}}{\sum_{l=1}^{n/m}\mu_l} + \frac{\gamma}{1-\gamma}\boldsymbol{\nu}^T \mathbf{r}_{(n)}\right) \\
&= \sigma^2\left(\sum_{h=1}^{n}\left[\left(\frac{\sum_{l=1}^{n/m}\mu_l p_{lh}}{\sum_{l=1}^{n/m}\mu_l} + \frac{\gamma}{1-\gamma}\nu_h\right) r(s^h_{ik})\right]\right).
\end{aligned}$$

Since $r(s^l_{ik})$ can be assumed to be independent and identically distributed (i.i.d.), we have

$$\begin{aligned}
&\sigma^2(J^*_{(n)}(s^{1\,\prime}_{ik})) \\
&= \sum_{h=1}^{n}\left[\frac{\sum_{l=1}^{n/m}\mu_l p_{lh}}{\sum_{l=1}^{n/m}\mu_l} + \frac{\gamma}{1-\gamma}\nu_h\right]^2 \sigma^2(r(s^h_{ik})) \\
&= \sum_{h=1}^{n}\left[\frac{c_h}{c} + \frac{\gamma}{1-\gamma}\nu_h\right]^2 \cdot \frac{1}{12}\left(\frac{L_i}{n}\right)^2
\end{aligned} \tag{7}$$

where $c = \sum_{l=1}^{n/m}\mu_l$ and $c_h = \sum_{l=1}^{n/m}\mu_l p_{lh}$.

Alternatively, $J^*_{(m)}(s^{1\,\prime}_{ik})$ can be obtained by $J^*_{(m)}(s^{1\,\prime}_{ik}) = \mathbf{p}'^T_1 \mathbf{r}_{(m)} + \gamma_i/(1-\gamma_i)\boldsymbol{\nu}'^T \mathbf{r}_{(m)}$. Using (5), p'_{1l} can be rewritten as $p'_{1l} = 1/c\sum_{h=h_{l-1}}^{h_l} c_h$, where $h_{l-1} = (n/m)(l-1)+1$ and $h_l = (n/m)l$. Hence, $\sigma^2(J^*_{(m)}(s^{1\,\prime}_{ik}))$ can be expressed as

$$\begin{aligned}
&\sigma^2(J^*_{(m)}(s^{1\,\prime}_{ik})) \\
&= \sigma^2\left(\sum_{l=1}^{m}\left(\frac{1}{c}\sum_{h=h_{l-1}}^{h_l} c_h + \frac{\gamma}{1-\gamma}\nu'_l\right) r(s^{l\,\prime}_{ik})\right) \\
&= \sum_{l=1}^{m}\left(\frac{1}{c}\sum_{h=h_{l-1}}^{h_l} c_h + \frac{\gamma}{1-\gamma}\nu'_l\right)^2 \cdot \frac{1}{12}\left(\frac{L_i}{m}\right)^2 \\
&= \sum_{l=1}^{m}\left(\sum_{h=h_{l-1}}^{h_l}\left(\frac{c_h}{c} + \frac{\gamma}{1-\gamma}\nu_h\right)\right)^2 \cdot \frac{1}{12}\left(\frac{L_i}{m}\right)^2
\end{aligned}$$

since $\nu'_l = \sum_{h=(n/m)(l-1)+1}^{(n/m)l}\nu_h$. Note that c_h/c and $\gamma_i/(1-\gamma_i)\nu_h$ are non-negative. Thus,

$$\sum_{h=h_{l-1}}^{h_l}\left(\frac{c_h}{c} + \frac{\gamma}{1-\gamma}\nu_h\right)^2 \le \left[\sum_{h=h_{l-1}}^{h_l}\left(\frac{c_h}{c} + \frac{\gamma}{1-\gamma}\nu_h\right)\right]^2.$$

Moreover, Cauchy–Schwarz inequality gives

$$\left[\sum_{h=h_{l-1}}^{h_l}\left(\frac{c_h}{c} + \frac{\gamma}{1-\gamma}\nu_h\right)\right]^2 \le \frac{n}{m}\sum_{h=h_{l-1}}^{h_l}\left(\frac{c_h}{c} + \frac{\gamma}{1-\gamma}\nu_h\right)^2.$$

Therefore, we have

$$\left(\frac{n}{m}\right)^2 \le \frac{\sigma^2(J^*_{(m)}(s^{1\,\prime}_{ik}))}{\sigma^2(J^*_{(n)}(s^{1\,\prime}_{ik}))} \le \left(\frac{n}{m}\right)^3. \tag{8}$$

■

TABLE I
PERFORMANCES ACHIEVED BASED ON MYOPIC (TFT) AND FORESIGHTED RESOURCE RECIPROCATION STRATEGIES

	Avg. Downloading Time	Avg. Video Quality
Myopic (TFT)	205.59 [sec]	34.05 [dB]
Foresighted	149.22 [sec]	35.44 [dB]

Proposition 1 implies that, in general, if more state descriptions are used, more accurate CDERs can be computed. Specifically, it provides upper and lower bounds for the accuracy improvement. This result can also be used to determine the minimum number of state descriptions required to achieve a tolerable CDER accuracy δ, as presented in Corollary 2.

Corollary 2: Suppose that a peer i with m state descriptions currently achieves a CDER accuracy σ_m by modeling the resource reciprocation of peer k. Then, a tolerable CDER accuracy δ for peer i can be achieved using at least $\lceil m(\sigma_m/\delta)^{2/3}\rceil$ state descriptions.

Proof: In (8), by setting $\sigma(J^*_{(m)}(s^{1\,\prime}_{ik})) = \sigma_m$ and $\sigma(J^*_{(n)}(s^{1\,\prime}_{ik})) = \delta$, it can be directly shown that $n \geq m(\sigma_m/\delta)^{2/3}$, which implies that the minimum number of state descriptions required to achieve δ is $\lceil m(\sigma_m/\delta)^{2/3}\rceil$. ■

Therefore, from Corollary 2, we can conclude that a peer that achieves a CDER accuracy σ_m using m state descriptions for its resource reciprocation strategy should use at least $\lceil m(\sigma_m/\delta)^{2/3}\rceil$ state descriptions to improve its CDER accuracy to δ.

IV. SIMULATION RESULTS

First, we quantify how much the proposed foresighted resource reciprocation strategies improve each peer's performance by comparing it with existing solutions such as TFT strategy in BitTorrent. The simulation results presented in Table I show the average time required for a peer i to complete its downloads (a general file with size of 5 Mbytes) and the quality measured in Peak Signal to Noise Ratio (PSNR) that a peer achieves by downloading *Foreman* sequences based on myopic (i.e., TFT) and the foresighted strategies. The sequence is at CIF (352×288) resolution, 30 frames/s, and encoded using the H.264/AVC encoder. The maximum available upload bandwidth L_k of associated peer k ($k \in C_i$) is 250 Kbps. As discussed in Section II, the foresighted resource reciprocation strategies maximize the CDERs of peers. This eventually leads the peers to download a file in a shorter average time or to achieve better video quality, as shown in Table I.

Since the peers can determine the number of their state descriptions by considering the accuracy of the CDERs, we next focus on quantifying the accuracy of CDERs based on different numbers of state descriptions. Simulation results are presented in Fig. 2. Note that the CDER of a peer in the results represents the cumulative discounted expected download rates from its associated peer, and the CDER accuracy σ implies a range of download rate variation. In this illustrative example, the available bandwidth L_k of associated peer k is 300 Kbps. Fig. 2 clearly shows that the accuracy of CDERs improves as the number of state descriptions increases, which is discussed in Proposition 1. Moreover, given a number of state descriptions, it is observed that a higher γ_i, which considers longer future rewards as shown in (1), induces a higher σ, implying an inaccurately computed CDER. Therefore, it can be generally concluded that a peer i with a higher γ_i requires more state descriptions than a peer j with a lower γ_j to achieve a tolerable CDER accuracy δ. For example, in Fig. 2, a peer with $\gamma_i = 0.1, 0.3, 0.5, 0.7$ and 0.9

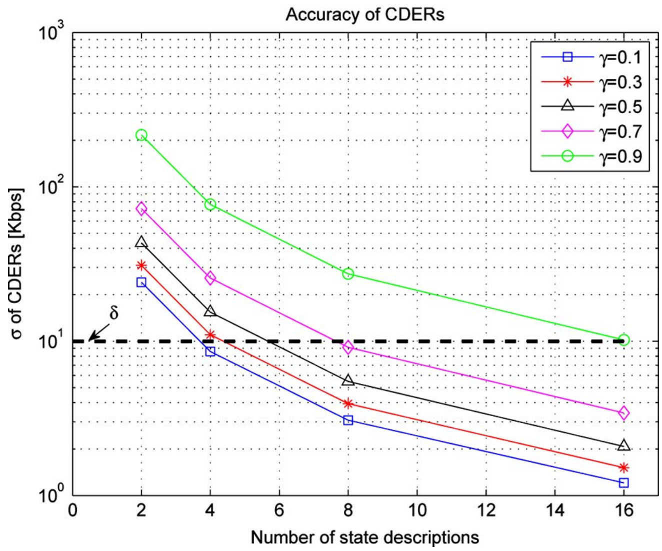

Fig. 2. CDER accuracy for different numbers of state descriptions and γ_i.

requires more than 4, 8, 8, and 16 state descriptions, respectively, to achieve $\delta = 10\,\mathrm{Kbps}$ in this resource reciprocation. Inversely, if the number of state descriptions is fixed, value of γ_i can be determined such that it satisfies a tolerable CDER accuracy δ. In this example, if the number of state descriptions is fixed as 8, $\gamma_i \leq 0.7$ needs to be selected, to achieve $\delta = 10\,\mathrm{Kbps}$.

We can also verify the result presented in Corollary 2. As an illustration, we consider a peer with $\gamma_i = 0.1$ and 2 state descriptions, which achieves $\sigma_2 = 25\,\mathrm{Kbps}$ as shown in Fig. 2. To achieve $\delta = 10\,\mathrm{Kbps}$, the minimum number of required state descriptions can be analytically computed based on the result in Corollary 2, which is $\lceil 3.68\rceil = 4$. Thus, using at least 4 state descriptions, the peer can achieve $\delta = 10\,\mathrm{Kbps}$, which can be verified in Fig. 2.

V. CONCLUSION

We analyze and quantify how the bounded rationality of peers affects their long-term utility and behavior. We analytically quantify how the number of state descriptions affects their complexity, memory requirements and ability to model the resource reciprocation of other peers, and, in turn, their own performance.

REFERENCES

[1] X. Zhang, J. Liu, B. Li, and T. S. P. Yum, "CoolStreaming/DONet: A data-driven overlay network for efficient live media streaming," *Proc. INFOCOM '05*, 2005.

[2] A. Legout, N. Liogkas, E. Kohler, and L. Zhang, "Clustering and sharing incentives in BitTorrent systems," *SIGMETRICS Perform. Eval. Rev.*, vol. 35, no. 1, pp. 301–312, 2007.

[3] K. G. Anagnostakis and M. B. Greenwald, "Exchange-based incentive mechanisms for peer-to-peer file sharingn," *Proc. ICDCS '04*, pp. 524–533, 2004.

[4] Q. Lian, Y. Peng, M. Yang, Z. Zhang, Y. Dai, and X. Li, "Robust incentives via multi-level tit-for-tat," *Proc. IPTPS '06*, 2006.

[5] M. Feldman, K. Lai, I. Stoica, and J. Chuang, "Robust incentive techniques for peer-to-peer networks," *Proc. ICEC '04*, pp. 102–111, 2004.

[6] H. Park and M. van der Schaar, "A framework for foresighted resource reciprocation in P2P networks," *IEEE Trans. Multimedia*, vol. 11, pp. 101–116, Jan. 2009.

[7] M. Piatek, T. Isdal, T. Anderson, A. Krishnamurthy, and A. Venkataramani, "Do incentives build robustness in BitTorrent?," *Proc. NSDI*, Apr. 2007.

[8] H. A. Simon, "A behavioral model of rational choice," *Quart. J. Econ.*, vol. 69, no. 1, pp. 99–118, Feb. 1955.

[9] M. L. Littman, T. L. Dean, and L. P. Kaelbling, "On the complexity of solving Markov decision problems," *Proc. UAI '95*, May 1995.