

Optimal Policies for Control of Peers in Online Multimedia Services

Young Myoung Ko, Jean-François Chamberland, and Natarajan Gautam

Abstract—In this paper we consider a distributed peer-based system with a centralized controller responsible for managing the peers. For this system customers request large volumes of information such as video clips which instead of retrieving from a centralized repository of a parent organization are obtained from peers that possess the clips. Peers act as servers only for a short duration and therefore the parent organization (i.e. centralized controller) would need to add new peer servers from time to time. This centralized “admission” control of deciding whether or not to admit a customer with a video clip as a peer based on the system state (number of waiting requests and number of existing peers) is the crux of this research. The problem can be posed as a discrete stochastic optimal control and is formulated using a Markov decision process approach with infinite horizon and discounted cost/reward. We show that a stationary threshold policy in terms of the state of the system is optimal. In other words the optimal decision whether or not to accept a customer as a peer server is characterized by a switching curve. In typical Markov decision processes, it is extremely difficult to derive an analytical expression for the switching curve. However, using an asymptotic analysis, by suitably scaling time and states taking fluid limits, we show how this can be done for our problem. In addition, the asymptotic analysis can also be used to show that the switching curve is independent of the model parameters such as customer arrival rate, downloading times and peer-server lifetimes. Several numerical results are presented to support the analytical claims based on asymptotic analysis.

I. INTRODUCTION

The demand for digital multimedia contents and delivery on the World Wide Web (WWW) has been growing rapidly. Increasing volume of contents implies service providers require large amount of network resources (e.g. network bandwidth, servers, storage, etc.). Until now, since the demand of digital multimedia contents has been limited to small music files and the advance of technology has made the resource costs cheaper, a service provider could provide services without imposing significant delay to customers. However, the demand for digital media is now moving to video files (e.g. movies, music videos, online lectures, user created contents, etc.). These kinds of contents require significantly more resources than what simple music files require. In addition, they may cause a service provider to suffer from maintaining adequate quality of service (QoS)

Young Myoung Ko is with Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX 77843-3131, USA ymko@tamu.edu

Jean-François Chamberland is with Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128, USA chmbrlnd@ece.tamu.edu

Natarajan Gautam is with Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX 77843-3131, USA gautam@tamu.edu

for customers. To address this problem, peer-to-peer (P2P) architecture is a viable alternative for multimedia companies to “outsource” service as it can cover bursts of demands and can offer a scalable capacity as the number of peers increases. Therefore user requests for digital multimedia contents can be satisfied by peers that have downloaded them in the past. The objective of this paper is to devise a control policy for the service provider to manage the peers with multimedia contents. The uniqueness of the scenario considered in this research are:

- *Content exchange* – Unlike free P2P networks (e.g. BitTorrent, eDonkey, etc.) which are based on a give-and-take philosophy, here customers pay for contents and all customers do not become peer servers.
- *Rewarding peers* – Peer servers are paid based on the time they are available for service as opposed to actual time spent serving or amount of contents; rewarding by availability encourages customers to join as peer servers more willingly.
- *QoS cost* – The discrete stochastic optimal control literature frequently uses holding cost for waiting customers. Here, however, a modified holding cost that we call QoS cost is considered.
- *Finite lifetime* – Peer servers have finite lifetime and can choose to become unavailable at any time (by moving, deleting content, network malfunctioning, etc.). In many P2P architectures it is assumed that peers are always available.

Considering the above issues, the objective of our research is to devise a control policy for our P2P architecture to determine when to allow a user to become a peer server (based on the number of users in the system and the number of peer servers) by trading off QoS cost against rewarding peers. Most of the recent research on P2P networks focuses on modeling and performance analysis (Ge et al [1], Clévenot and Nain [2], Qiu and Srikant [3]), or optimal peer search and selection (Adler et al [4]). However control of P2P networks is relatively unexplored. To the best of our knowledge, only Tewari and Kleinrock [5] and, Cohen and Shenker [6] deal with management strategies in P2P networks. However, their main objective is to improve the effectiveness of content search. Furthermore, their concerns are not P2P networks managed by a centralized commercial company but rather free P2P networks. In addition, the research uses a deterministic analysis as opposed to the stochastic control considered here.

To address our research objective, we formulate our prob-

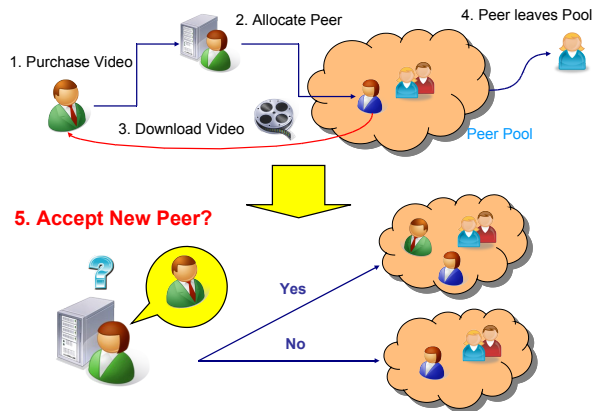


Fig. 1. Problem illustration

lem as a Markov Decision Process (MDP). To obtain the optimal policy we use techniques such as truncating state space and perform value-iteration with uniformization (see Bertsekas [7]). We make three interesting observations: i) a stationary threshold policy is optimal, ii) the optimal policy is invariant to input parameters and is only affected by the cost function, and iii) asymptotic closed form expression for the switching curve is obtained. In fact, the optimality of a stationary threshold policy is not surprising since many other admission control problems display such behavior. However, the parameter-invariance and closed form expression of the switching curve are unique. We explain these two findings through a fluid approximation by scaling time and states. Fluid approximation is commonly used for asymptotic behavior of decision process since it is relatively tractable analytically (Altman et al [8] and Savin et al [9]).

The rest of the paper is organized as follows. In Section II, we describe our problem and formulate it as an MDP. Although our model has an infinite state space, by truncating the state space, we illustrate the stationary threshold policy being optimal in Section III. In Section IV, we explain how the optimal policy looks in the asymptotic regime using fluid approximation. We obtain the asymptotic optimal policy as a closed-form algebraic expression and it also explains the parameter-invariance. Finally, in Section V, we make some concluding remarks and suggest directions for future work.

II. MODEL AND ANALYSIS

A. Problem statement

Fig. 1 is a simplified illustration of our problem. We consider an online entertainment company that sells digital video contents via the WWW. The company has a main server that stores video contents. Customers access the main server and purchase contents through the company's web site. The company operates P2P networks consisting of peers who purchased that contents before and are given the authorization for delivering contents to new customers. After a new customer arrives, he/she waits for a free "server" (an appropriate peer or the main server itself) to be assigned.

After a download is complete, the company decides whether or not the customer that just finished downloading should be admitted as a new peer. Each peer server stays available for service for a limited time. The peers in P2P network are rewarded by the company based on the time they spend as a peer. The P2P network grows whenever the company accepts a new peer and shrinks whenever a peer leaves the system. The objective of this paper is to find an optimal policy for the company to decide when it has to accept or to reject a new peer. Keeping our objective in mind, we formulate this problem as an MDP.

B. MDP formulation

1) *System states*: We define the state of system $\{\bar{X}(t)\}$ as a two-dimensional stochastic process $\bar{X}(t) = (x(t), y(t))$ where $x(t)$ is the number of customers and $y(t)$ is the number of peers in the P2P pool (or network) at time t . Before formulating the MDP we need to identify the type of events to decide if a control action is appropriate. In our problem, three types of events can occur: customer arrival, service completion, and peer's departure. Since a control action is taken only when a service completion occurs, we need an indicator that tells us the event is a service completion. To reflect this, we amend the system state slightly as follows. Assume that all peers are managed in a peer pool. When a service completes, the customer that just finished getting served is sent to a waiting space (not to the peer pool). He/she should stay there until next event occurs regardless of the decision (i.e. accepted or rejected). If accepted, he/she starts service in a waiting space and at the time of next event is sent to the peer pool. If rejected, he/she stays at a waiting space and leaves the system when next event occurs. With this adjustment, define $X(t) = (x(t), y(t), z(t))$ as the state of the system where $z(t)$ is the number of customers in the waiting space, and $x(t)$ and $y(t)$ are as defined earlier. Note that since the customer in the waiting space stays for only one inter-event time, $z(t)$ can only take the value 0 or 1. For the next step, define the control variable $U(\cdot)$. The variable $U(\cdot)$ depends on the value of $z(t)$ because a decision is taken about adding a peer only when a service is completed; if any other event occurs, no decision is required and the company doesn't do anything. Therefore,

$$U(z) = \begin{cases} \{0\} & \text{if } z = 0 \\ \{0, 1\} & \text{if } z = 1 \end{cases} \quad (1)$$

In (1), when $z = 1$, the control value 1 means accepting a new peer and 0 means rejecting a new peer.

Now we have the states and control space of the system. The next thing we can do is defining the transition probability for the MDP formulation.

2) *Transition probability*: We assume that customers arrive to the system according to a Poisson process with rate λ and the service time of each customer is exponentially distributed with mean $1/\mu$ in both the main server and peers. If a customer becomes a peer, he or she stays as a peer for a random amount of time that follows an exponential distribution with mean $1/\theta$. With these assumptions, let T denote the time between decision epochs and $Q(T <$

$\tau, (x', y', z')|(x, y, z), u$ denote the transition probability from state (x, y, z) to (x', y', z') given control u . Then, for $\tau > 0$, Q is defined as follows;

$$Q(T < \tau, (x+1, y+zu, 0)|(x, y, z), u) = \frac{\lambda}{\beta(x, y+zu)}(1 - e^{-\beta(x, y+zu)\tau}), \quad (2)$$

$$Q(T < \tau, (x-1, y+zu, 1)|(x, y, z), u) = \frac{\min(x, y+zu)\mu}{\beta(x, y+zu)}(1 - e^{-\beta(x, y+zu)\tau}), \quad (3)$$

$$Q(T < \tau, (x, y+zu-1, 0)|(x, y, z), u) = \frac{(y+zu)\theta}{\beta(x, y+zu)}(1 - e^{-\beta(x, y+zu)\tau}), \quad (4)$$

where $\beta(x, y) = \lambda + \min(x, y)\mu + y\theta$.

3) *Cost function*: In the previous sections, we defined the states and obtained the transition probabilities. Now we move to the final step of the MDP formulation, namely, defining the cost functions. The objective of our study is to find the optimal peer admission policy that minimizes the cost. Cost is incurred in two ways: service delay (related to the number of customers in queues) which we will call QoS cost and maintenance of peers (related to the number of peers). Therefore based on these factors the cost per unit time $g(\cdot, \cdot)$ is defined as follows:

$$g(X, u) = c\left(\frac{(x - (y + zu))^+}{y + zu}\right) + q(y + zu), \quad (5)$$

where $c(\cdot)$ and $q(\cdot)$ are functions representing the QoS cost of customers and the reward cost of peers respectively, $(x)^+ = \max(x, 0)$, and $X = (x, y, z)$. In the function $c(\cdot)$, the numerator is the number of customers waiting in queues to begin service and the denominator denotes the number of queues (i.e. peers). It uses the instantaneous queue length per peer for the argument of $c(\cdot)$ instead of the total number of customers in the queue or system. It is natural to think like this since the cost incurred when there are 10 peers and 10 waiting customers in queues is different from that incurred when 2 peers and 10 waiting customers in queues. From (2)-(4), we know that the length between decision epochs (T) is random. Therefore, the optimal cost J^* is defined based on the description in Bertsekas [7]:

$$J^*(X(0)) = \min \lim_{n \rightarrow \infty} E \left\{ \int_0^{t_n} e^{-\gamma t} g(X(t), u(t)) dt \right\},$$

where, t_n is the time of n^{th} event and γ is the discounting factor. For continuous time models like this problem, uniformization technique is usually considered to modify the problem into a discrete time problem. However, the uniformization technique cannot be applied to this problem directly since the transition rate (β) is not bounded. To resolve this, we look back to the system. The company manages peers and may have the capacity in managing peers. Since the capacity is naturally finite, it is acceptable to restrict the maximum number of peers if it is defined reasonably.

C. Model adjustment

Let M denote the maximum number of peers that the company allows. Then M should be chosen carefully since a small M might make the system unstable (i.e. $M\mu < \lambda$).

Assume that the system is stable with M such that $M\mu > \lambda$. Then, the transition rate is bounded by $\lambda + M(\mu + \theta)$. Using the method described in Bertsekas [7], the new uniformized transition probabilities (P) are obtained as follows:

$$P((x+1, y+zu, 0)|(x, y, z), u) = \frac{\lambda}{\lambda + M(\mu + \theta)},$$

$$P((x-1, y+zu, 1)|(x, y, z), u) = \frac{\min(x, y+zu)\mu}{\lambda + M(\mu + \theta)},$$

$$P((x, y+zu-1, 0)|(x, y, z), u) = \frac{(y+zu)\theta}{\lambda + M(\mu + \theta)},$$

$$P((x, y, z)|(x, y, z), u) = \frac{(M - \min(x, y+zu))\mu + (M - y - zu)\theta}{\lambda + M(\mu + \theta)}.$$

Since the cost function satisfies Assumption P in Bertsekas [7], there exists a stationary policy and the optimal cost function J^* satisfies the Bellman equation, i.e.,

$$\begin{aligned} J^*(x, y, 0) = & R \left(g((x, y, 0), 0) + \lambda J^*(x+1, y, 0) \right. \\ & + \min(x, y)\mu J^*(x-1, y, 1) \\ & + y\theta J^*(x, y-1, 0) \\ & + \left\{ (M - \min(x, y))\mu \right. \\ & \left. + (M - y)\theta \right\} J^*(x, y, 0) \left. \right), \end{aligned} \quad (6)$$

$$\begin{aligned} J^*(x, y, 1) = & \min_{u \in \{0,1\}} R \left(g((x, y, 1), u) + \lambda J^*(x+1, y+u, 0) \right. \\ & + \min(x, y+u)\mu J^*(x-1, y+u, 1) \\ & + (y+u)\theta J^*(x, y+u-1, 0) \\ & + \left\{ (M - \min(x, y+u))\mu \right. \\ & \left. + (M - y - u)\theta \right\} J^*(x, y, 1) \left. \right), \end{aligned} \quad (7)$$

where $R = 1/(\lambda + M(\mu + \theta) + \gamma)$ and γ is a discount factor.

There are several ways to obtain the optimal cost and policy. In this paper, we use the value iteration method. In the following section, we will show some interesting results using numerical examples.

III. VALUE ITERATION RESULTS

In the previous section, in order to utilize the uniformization technique, we assume that there is a limit on the maximum number of peers (M). First, we conduct value iteration for different M 's to study the effect of M . Then, we move to the effects of the parameters (i.e. λ, μ, θ) and effects of the different cost functions. The optimal policy obtained by solving the Bellman equation using value iteration is a stationary threshold type policy called "switching curve". To prove the optimality of threshold policy mathematically, some conditions (see Porteus [10]) need to be shown as being satisfied. Due to space restrictions, we will provide the detailed mathematical proof only in an extended version of this paper; we will also provide the plots of the actual costs obtained from value iteration in the extended version. Fig. 2 shows a typical shape of a switching curve. Horizontal and vertical axes represent the number of customers and the

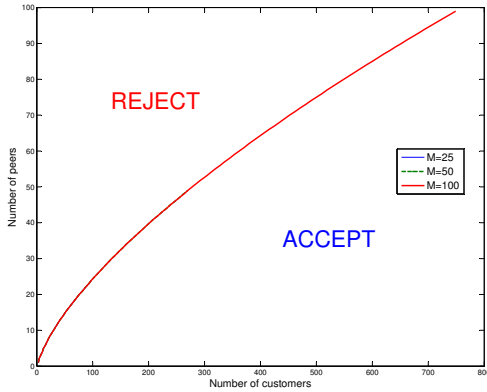


Fig. 2. The effects of maximum number of peers ($M = 25$ vs $M = 50$ vs $M = 100$)

number of peers in the system respectively. If the system state lies to the right side of the curve, the optimal control action is to accept the customer that completed downloading to become a peer and join the pool.

A. Optimality of the policy obtained from the finite state space model

We consider three examples that have the same parameters $\lambda = 50$, $\mu = 3$, $\theta = 1$. However the maximum number of peers (M) is selected as 25, 50, and 100. Fig. 2 shows the switching curves for our three examples. As seen in Fig. 2, the optimal policies are exactly overlapped (except that a close observation would reveal that for the case of $M = 25$ the maximum number of peers stop at 25). The boundary of each example is not plotted since on the boundary there is no choice but to reject a new peer. When we perform value iteration for several other examples, exactly the same results are obtained. Therefore, we can conclude that M does not affect the optimal policy.

This phenomenon can be explained using the Bellman equation defined in (6) and (7). By adjusting the left and right hand side of (6) and (7), we obtain

$$\left[\lambda + \left\{ \min(x, y + zu^*)\mu + (y + zu^*)\theta \right\} + \gamma \right] J^*(x, y, z) = g((x, y, z), u^*) + \lambda J^*(x + 1, y + zu^*, 0) + \min(x, y + zu^*)\mu J^*(x - 1, y + zu^*, 1) + (y + zu^*)\theta J^*(x, y + zu^* - 1, 0) \quad (8)$$

where $g(\cdot, \cdot)$ is as given in (5) and u^* is the action that minimizes the right hand side of (6) and (7).

As seen in (8), we find that M does not affect the recursion relation of J^* . If all other parameters are same, the optimal policy would also be same for all states within the boundary.

B. Effect of input parameters

Now we move our attention to the effect of parameters such as λ , μ and θ . First, we observe the change of the optimal policy when λ changes. In fact, we conjecture that the threshold value becomes small as λ increases. However, the observation shows different results. Fig. 3 shows the

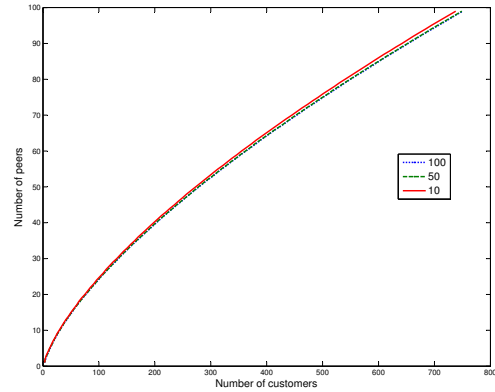


Fig. 3. The effects of arrival rates ($\lambda = 100$ vs $\lambda = 50$ vs $\lambda = 10$)

threshold curves when λ has values 10, 50, and 100. As shown in Fig. 3, the changes on the threshold values are very small although the arrival rate becomes ten times bigger. To see whether this phenomenon occurs for other parameters, we observe the results by changing other parameters (μ and θ). However, the changes of μ and θ show a negligible effect on the threshold values of the optimal policy (due to space limitation, we do not illustrate them). The mathematical explanation of this phenomenon will be provided in Section IV.

So far we have shown that the parameters (M , λ , μ , and θ) have no significant effect on the optimal policy. The question that comes to mind immediately is then what affects the optimal policy? We address that next.

C. The effects of the cost functions

As described earlier, the cost function in this problem consists of two parts: QoS cost of customers and reward cost of peers. Intuitively, the latter can be thought as a linear function of the number of peers. However, the QoS cost of customers should be defined carefully since it reflects how customer delay is treated. If customer delay is critical, the QoS cost should be increasing rapidly as the number of customers increases and vice versa. In that light, we observe the changes of the optimal policy when the cost functions change. Three different QoS costs $c_1((x, y, z), u)$, $c_2((x, y, z), u)$, and $c_3((x, y, z), u)$ are considered for comparison and for which the optimal policy is plotted in Fig. 4. They are defined as follows:

$$\text{Linear: } c_1((x, y, z), u) = \frac{(x - y - zu)^+}{y + zu} \quad (9)$$

$$\text{Quadratic: } c_2((x, y, z), u) = \left(\frac{(x - y - zu)^+}{y + zu} \right)^2 \quad (10)$$

$$\text{Cubic: } c_3((x, y, z), u) = \left(\frac{(x - y - zu)^+}{y + zu} \right)^3 \quad (11)$$

Going from $c_1(\cdot, \cdot)$ to $c_3(\cdot, \cdot)$, the system becomes more time critical. Fig. 4 shows intuitive results. If the customers' delay is crucial, the system accepts a new peer earlier to reduce

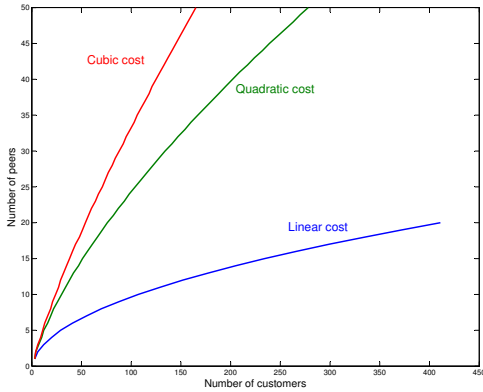


Fig. 4. The effects of the cost function

the delay. On the other hands, if the customers' delay is less crucial, the system accepts a new peer later to reduce the maintenance cost of peers.

IV. FLUID APPROXIMATION

The next step is to perform an asymptotic analysis in order to see if we could describe the switching curve as an analytical expression. As a byproduct, the analysis also shows in the limit the switching curve is independent of model input parameters. Qiu and Srikant [3] use fluid and diffusion approximations to model the behavior of P2P networks since it is not tractable to model it using standard queueing methods even if the system has Markovian properties. Usually, these approximations scale both in time and state space, and get a limit process that is mathematically tractable using functional strong law of large numbers (FLLN) and functional central limit theorem (FCLT); for the details of FLLN and FCLT, see Whitt [11]. Inspired by their work, we apply the scaling of both time and state space to our Bellman equation (7). In order to apply scaling, we adjust our model slightly. Let $X(t) = (x(t), y(t), z(t))$ denote the state of the system at time t . For fluid approximation, the new state space S' is set as follows:

$$S' = A \times A \setminus \{0\} \times \{0, \epsilon\}.$$

where, $A = \{0, \epsilon, 2\epsilon, \dots\}$ and $\epsilon > 0$ is a scaling factor. Note that since rational numbers are countable and dense in $[0, \infty)$, as $\epsilon \rightarrow 0$, $S' \rightarrow S$, where

$$S = \mathbf{R}^+ \times \mathbf{R}^+ \setminus \{0\} \times \{0\}.$$

Then, we can rewrite our Bellman equations defined in (7) as follows:

$$\begin{aligned} J^*(x, y, \epsilon) = & \min_{u \in \{0,1\}} R \left(\frac{1}{\epsilon} g((x, y, \epsilon), u) \right. \\ & + \frac{\lambda}{\epsilon} J^*(x + \epsilon, y + u\epsilon, 0) \\ & + \frac{\min(x, y + u\epsilon)\mu}{\epsilon} J^*(x - \epsilon, y + u\epsilon, \epsilon) \\ & \left. + \frac{(y + u\epsilon)\theta}{\epsilon} J^*(x, y + u\epsilon - \epsilon, 0) \right) \end{aligned}$$

$$\begin{aligned} & + \frac{1}{\epsilon} \left\{ (M - \min(x, y + u\epsilon))\mu \right. \\ & \left. + (M - y - u\epsilon)\theta \right\} J^*(x, y, \epsilon), \end{aligned} \quad (12)$$

where $R = \epsilon / (\lambda + M(\mu + \theta) + \gamma)$ and γ is a discount factor.

The value ϵ is a scaling constant which in spirit is identical to $1/\eta$ used for scaling time and state space in fluid approximations in Massey [12]. With (12), we could think that on the boundary of the threshold, whichever action we choose, the optimal cost is the same. Therefore

$$J(x, y, \epsilon)_{u=1} = J(x, y, \epsilon)_{u=0}. \quad (13)$$

Combining (12) and (13), we obtain the following relation:

$$\begin{aligned} & - \left\{ g((x, y, \epsilon), 1) - g((x, y, \epsilon), 0) \right\} \\ & = \lambda \left\{ J^*(x + \epsilon, y + \epsilon, 0) - J^*(x + \epsilon, y, 0) \right\} \\ & + \mu \left\{ \min(x, y + \epsilon) J^*(x - \epsilon, y + \epsilon, \epsilon) \right. \\ & \left. - \min(x, y) J^*(x - \epsilon, y, \epsilon) \right\} \\ & + \theta \left\{ (y + \epsilon) J^*(x, y, 0) - y J^*(x, y - \epsilon, 0) \right\} \\ & - \left\{ (\min(x, y + \epsilon) - \min(x, y))\mu + \epsilon\theta \right\} J^*(x, y, \epsilon). \end{aligned} \quad (14)$$

Note that scaling for the time space disappears in (14) as $\epsilon \rightarrow 0$. For $x > y$ and $y + \epsilon \leq x$, we rewrite (14) as follows:

$$\begin{aligned} & - \left\{ g((x, y, \epsilon), 1) - g((x, y, \epsilon), 0) \right\} \\ & = \lambda \left\{ J^*(x + \epsilon, y + \epsilon, 0) - J^*(x + \epsilon, y, 0) \right\} \\ & + \mu \left\{ (y + \epsilon) J^*(x - \epsilon, y + \epsilon, \epsilon) - y J^*(x - \epsilon, y, \epsilon) \right\} \\ & + \theta \left\{ (y + \epsilon) J^*(x, y, 0) - y J^*(x, y - \epsilon, 0) \right\} \\ & - (\mu + \theta) J^*(x, y, \epsilon). \end{aligned} \quad (15)$$

Dividing by ϵ in both sides of (15) and taking limit (i.e. $\epsilon \rightarrow 0$), we can obtain the following form:

$$\frac{\partial J^*(x, y, 0)}{\partial y} = \frac{-g^*(x, y)}{\lambda + (\mu - \theta)y}, \quad (16)$$

where $g^*(x, y) = \lim_{\epsilon \rightarrow 0} \frac{g((x, y, \epsilon), 1) - g((x, y, \epsilon), 0)}{\epsilon}$.

Likewise, assuming that $\lambda + \mu x - \theta y \neq 0$, for $x \leq y$, we obtain the following:

$$\frac{\partial J^*(x, y, 0)}{\partial y} = \frac{-g^*(x, y)}{\lambda + \mu x - \theta y}. \quad (17)$$

Since $\frac{\partial J^*(x, y, 0)}{\partial y} = 0$ on the boundary of threshold, by setting the left-hand side of (16) and (17) to zero, we can obtain the formula for threshold as the solution of the following equation:

$$g^*(x, y) = 0. \quad (18)$$

Note that the solution of (18) is independent of the model parameters (λ, μ , and θ). Therefore, we can explain the invariance on the model parameters observed in Section III with (18).

Now, we verify how our threshold formula (18) works well against the results of value iteration in Section III. Fig. 5 and 6 show the plots of the threshold values obtained from value

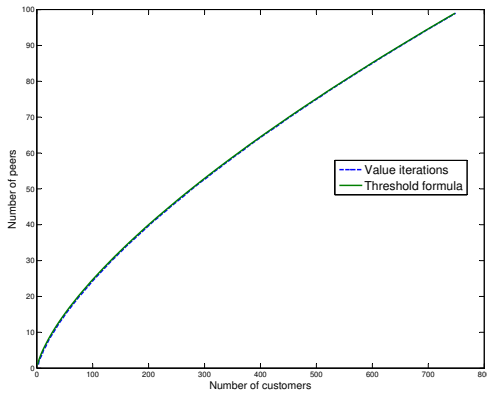


Fig. 5. Value iteration vs threshold formula when $g = \left(\frac{(x-y)^+}{y}\right)^2 + y$

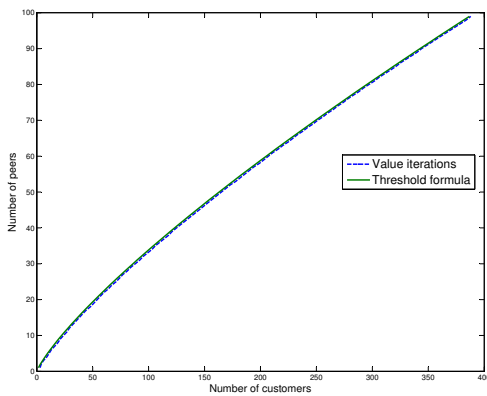


Fig. 6. Value iteration vs threshold formula when $g = \left(\frac{(x-y)^+}{y}\right)^3 + y$

iteration and threshold formulae corresponding to quadratic and cubic QoS cost functions (10) and (11) respectively; although we have a threshold formula $x = y^2$ for the linear QoS cost function (9), due to space restrictions we would not provide a figure of that case. For each case, we obtain the threshold formulas as follows:

$$x = \frac{y(1 + \sqrt{1 + 2y})}{2} \quad \text{for quadratic QoS cost}$$

$$x = \left\{ \frac{2}{3} + \frac{1}{6} \left(-8 + 36y + 12\sqrt{-4y + 9y^2} \right)^{\frac{1}{3}} + \frac{2}{3\sqrt{-8 + 36y + 12\sqrt{-4y + 9y^2}}} \right\} \cdot y$$

for cubic QoS cost. As seen in Fig. 5 and 6, the plot obtained by our threshold formula is fairly close to the plots obtained by value iteration when no scaling is done.

V. CONCLUSION

In this paper we formulate a discrete state and discrete action stochastic optimal control problem to manage peers

for online entertainment companies that carry multimedia files like videos. The control problem results in a Markov decision process and the optimal policy is obtained using value iteration method by truncating the state space suitably. Although the truncation is an approximation we show that the optimal policy does not change with the truncation limit. In fact the optimal policy is also practically invariant with respect to input parameters (such as arrival rate, service time and server lifetime parameters). Since the input parameters do not influence the threshold values of the optimal policy significantly and only the cost function affects them drastically, it is possible to conjecture that if the parameters are not constant (e.g. varying over time) the optimal policy might still be invariant. The most significant contribution of this research is a characterization of the switching curve obtained for the optimal threshold policy. We use fluid limits for this characterization and derive a formula for the threshold using an asymptotic analysis. We show that this asymptotic result is extremely close to the regular non-asymptotic case that can be numerically evaluated using value iteration. In addition, the fluid limits also can be used to show the invariance of the optimal policy to the input parameters.

There are several extensions to this research that can be considered in the future. Firstly, the peer-servers can go into an alternating on-off mode. Secondly, the video file itself could become obsolete after a period of time and thereby the optimal policy may not be stationary. Thirdly, one could consider relaxing some of the assumptions made in the model formulation.

REFERENCES

- [1] Z. Ge, D. R. Figueiredo, S. Jaiswal, J. Kurose, and D. Towsley, "Modeling peer-to-peer file sharing systems," in *Proc. IEEE INFOCOM'03*, 2003.
- [2] F. Clévenot and P. Nain, "A simple fluid model for the analysis of the squirrel peer-to-peer caching system," in *Proc. IEEE INFOCOM'04*, 2004.
- [3] D. Qiu and R. Srikant, "Modeling and performance analysis of bittorrent-like peer-to-peer networks," *SIGCOMM Comput. Commun. Rev.*, vol. 34, no. 4, pp. 367–378, October 2004. [Online]. Available: <http://portal.acm.org/citation.cfm?id=1030194.1015508>
- [4] M. Adler, R. Kumary, K. Rossz, D. Rubensteinx, T. Suel, and D. D. Yao, "Optimal peer selection for p2p downloading and streaming," in *Proc. IEEE INFOCOM'05*, 2005.
- [5] S. Tewari and L. Kleinrock, "Proportional replication in peer-to-peer networks," in *Proc. IEEE INFOCOM'06*, 2006.
- [6] E. Cohen and S. Shenker, "Replication strategies in unstructured peer-to-peer networks," in *Proc. SIGCOMM'02*, 2002.
- [7] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 2nd ed. Athena Scientific, 2001.
- [8] E. Altman, T. Jimenez, and G. Koole, "On optimal call admission control in a resource-sharing system," *IEEE Transactions on Communications*, vol. 49, no. 9, pp. 1659–1668, 2001.
- [9] S. V. Savin, M. A. Cohen, N. Gans, and Z. Katalan, "Capacity management in rental businesses with two customer bases," *Operations Research*, vol. 53, no. 4, p. 617.631, 2005.
- [10] E. Porteus, "Conditions for characterizing the structure of optimal strategies in infinite-horizon dynamic programs," *Journal of Optimization Theory and Applications*, vol. 36, no. 3, p. 419.432, mar 1982.
- [11] W. Whitt, *Stochastic Process Limits*, 1st ed. Springer, 2002.
- [12] W. A. Massey, "The analysis of queues with time-varying rates for telecommunication models," *Telecommunication Systems*, vol. 21, no. 2-4, pp. 173–204, dec 2002.