Applied Mathematics & Information Sciences
An International Journal

A Dynamic SGP Selection Algorithm based on Evolutionary Game for Hybrid P2P Streaming System

Jing Chen[1,2], *Yangjie Cao*[3], *Li Li*[4] *and Xiaoshe Dong*[1]

[1]School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China
[2]School of Information Engineering, Zhengzhou University, Zhengzhou 450052, China
[3]School of Software Technology, Zhengzhou University, Zhengzhou 450052, China
[4]College of Computer Science, Shanxi Normal University,Xi'an 710049, China

Received: 26 Oct. 2012, Revised: 10 Dec. 2012, Accepted: 21 Dec. 2012
Published online: 1 May 2013

Abstract: Due to high efficiency and good scalability, hierarchical hybrid P2P architecture has drawn more and more attentions in P2P video streaming applications recently. The problem about Super Group Peers (SGPs) selection, which is the key problem in hierarchical hybrid P2P architecture, is becoming highly challenging because super peers must be selected from a huge and dynamically changing network. In this paper, we propose a SGP selection game model based on evolutionary game framework and analyze its evolutionarily stable strategies in theory. Moreover, we propose a distributed Q-Learning algorithm, which has the ability to make the peers converge to the ESSs based on its own payoff history. Compared to the randomly super peer selection scheme in traditional P2P streaming systems, experiment results show that the proposed algorithm achieves better performance in terms of social welfare, average upload rates of SGPs, and keeps the upload capacity of the P2P streaming system increasing steadily with the number of peers increasing.

Keywords: P2P streaming system, Super group peer selection, Evolutionarily, Game theory, Evolutionarily Stable Strategy

1. Introduction

According to the characteristic of data structure, current P2P media streaming system deployed can be generally divided into three types: tree-push based topology, mesh-pull based topology and hierarchical hybrid topology.Among them, the hierarchical hybrid topology model, which divides the peers in P2P streaming systems into super nodes and ordinary nodes according to the performance differences of peers, has become the focus of recent researches. The hierarchical hybrid model integrates the advantages of tree-based topology and mesh-based topology. It allows the nodes with stronger ability to participate fully in the P2P streaming system to get higher scalability and stability. However, the super peers selection problem, which is the key problem in hierarchical hybrid P2P system, is highly challenging because in the peer-to-peer environments a large number of super peers must be selected from a huge and dynamically changing network in which neither the node characteristics nor the network topology are known a priori [1]. Furthermore, the super peers selection plays an important role in the efficiency and the Quality of Experience (QoE) of the whole P2P streaming system.

Currently, there are many research literatures about super peers selection problem [1–6]. Most of them, however, focus on the capacity of peers [1,4,5], such as memory space, network bandwidth, CPU cycles, etc. And few studies focus on the behavioral attribute of peers. Super group peers need to provide resources inquiry services and routing services for the other ordinary nodes within the group, and they also need to upload video streaming chunks for them. The cost of calculation and storage of super peers may lead to their system performance decline. To solve this problem, Virginia Lo et al. proposed the H_2O (Hierarchical 2-level Overlay) protocol [1] to select super group peers in unstructured P2P networks, which can take trust, secure paths, routing performance and some other factors into account. However the notification messages between super peers in H_2O are distributed through flooding mechanism, resulting in a lot of network overhead. Y.Jin et al. proposed a super peer selection method based on proxy trust in [7]. They used proxy trust to identify the

* Corresponding author e-mail: iejchen@zzu.edu.cn

behaviors of super peers. With this metric a peer can select its own trusted super peers and isolate malicious super peers in the system. The method proposed in [7] reduced the overhead of network flooding in [1] to some extent, but it required a priori knowledge in the decision on whether to trust a node or not. Game theory is a suitable mathematic tool to model the strategy interaction between two or more decision makers, and it has attracted more and more attention in computer science in recent years[8,9]. The peers in P2P video streaming system are rational and selfish, and they can make decisions according to their needs and the behavior of other peers. In this paper, we used evolutionary game theory framework to model the super peers selection issue in P2P streaming system, and the Evolutionarily Stable Strategy (ESS) is analyzed.

In traditional game theory, the Nash equilibrium is the most popular solution. It ensures that a player cannot improve its payoff if none of the other players in the game deviates from the solution. However, when the solution to a non-cooperative game has multiple Nash equilibriums, a refined solution is required. Evolutionary equilibrium, which is based on the theory of evolutionary game theory, provides such a refined solution, and it ensures stability (i.e., group of players will not change their chosen strategies over time). An evolutionary game can explicitly capture the dynamics of interaction among the players in a population. In an evolutionary game, a player can observe the behavior of other players, learn from the observations, and make the best strategy based on its knowledge. In addition, with replicator dynamics, the state of the game can be determined at a particular point in time, which is useful for investigating the trends of the strategies of the players while adapting their behavior to reach the solution.

The rest of the paper is organized as follows. In Section 2, we describe the system model and utility function. In Section 3, we model the SGP selection procedure by evolutionary game. We analyzes the SGP selection evolutionary game model by applying replicator dynamics equation in section 4 and proposed a distributed Q-Learning algorithm for SGP selection evolutionary game model to convergent to the ESS in section 5. In section 6 we show experiment results to evaluate the performance of the proposed ESS-SGP algorithm. Last we conclude the paper and propose future work in Section 7.

2. System Model and Utility Function

2.1. System Model

Peers in hierarchical hybrid P2P streaming system are divided into peer groups according to their geographical location [10] as shown in Fig. 1. We expect that the video streaming chunks of a certain channel can be shared within one peer group as much as possible to reduce unnecessary network traffic. There are a set of group peers who want to view a live video streaming channel simultaneously in the edge of the hierarchical hybrid P2P live streaming system.Every peer can choose to act as a Super Group Peer (SGP) or an Ordinary Peer (OP) in a group. If a peer choose to act as a SGP, then he/she not only need to act as a client to download video streaming chunks from the SGPs in other peer groups, but also need to act as a server to upload video streaming chunks for both the SGPs in other groups and the peers in the same group.

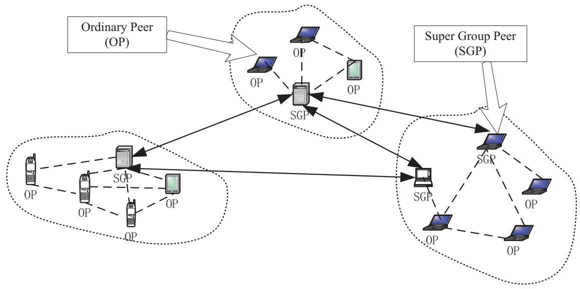

Figure 1 Peer groups at the edge of hierarchical hybrid P2P streaming system.

On the contrary, if a peer chooses to be an OP, he/she only needs to download or upload necessary streaming chunks in the same group. Assume that the upload and download bandwidth within a super group is larger than that cross groups. In such a case, peers tend to be an OP because of their selfish nature. However, from the other aspect, the OPs have a risk of receiving degraded streaming quality since there may not be sufficient SGPs to download streaming chunks from other groups.

2.2. Utility Function

In the hierarchical hybrid P2P streaming system, peers not only act as clients who download needed streaming chunks from other peers, but also act as servers to upload video streaming chunks for the other peers. So, while a peer can benefit downloading video streaming chunks from the other peers, he/she also causes a cost in uploading video streaming chunks for the other peers, where the cost may be the resource consumed on uploading chunks, such as bandwidth, memory, buffer size, and so on.

Assume that there are N peers within the group, among which, s peers are willing to act as SGPs to download video streaming chunks from the SGPs of other groups. Suppose that the download rates of the s SGPs are $r_1, r_2, \cdots, r_s$, then the total download rate of the group peers is $d_s = \sum_{i=1}^{s} r_i$.

As these s SGPs select peers outside the group for downloading streaming chunks randomly and independently, the download rate r_i's are random variables. According to [11], the Cumulative Distribution Function (CDF) of a peer's download bandwidth can be modeled as a linear function, which means that the Probability Density Function(PDF) of a peers download bandwidth can be viewed as a uniform distribution, which means r_i's are uniformly distributed.

Obviously, if the total download rate d_s is no less than the source rate r, then the group peers can obtain the effect of real-time streaming, and all the group peers can obtain a certain profit G. Otherwise there will be some delay, and in this case we assume the gain of the peer is 0. Therefore, given the total download rate d_s and the source rate r, if peer i chooses to be a SGP, then its utility function is presented by

$$\pi_{SGP,\ i}(s) = Pr(d_s \geq r)G - C_i, \forall s \in [1,N], \tag{1}$$

where C_i is the cost of peer i when he/she acts as a SGP, and $Pr(d_s \geq r)$ is the probability that the peer can obtain real-time streaming effect. If we assume that $r_1, r_2, \cdots, r_s$ are independent and identically distributed in $[r^L, r^U]$, and s is sufficiently large, then $Pr(d_s \geq r)$ can be approximately computed as $Pr(d_s \geq r) \approx Q(\frac{\hat{r}-s/2}{\sqrt{s/12}})$, where $\hat{r} = \frac{r-sr^L}{r^U-r^L}$ and $Q(x)$ is the Gaussian tail function $\int_x^\infty \frac{1}{\sqrt{2\pi}} \exp^{-\frac{x^2}{2}} dx$.

As the upload and download bandwidths within the group are large, it is assumed that the cost of uploading the streaming chunks to the other peers within the same group can be negligible. In this case, if a peer chooses to be an OP instead of a SGP, then there is no cost for him/her, and the utility function is:

$$\pi_{OP,\ i}(s) = \begin{cases} Pr(d_s \geq r)G & \text{, if } s \in [1,N-1]; \\ 0 & \text{, if } s = 0. \end{cases} \tag{2}$$

3. Super Group Peer Selection Evolutionary Game Model

We adopt the concept of Evolutionarily Stable Strategy (ESS) [12] to provide a robust equilibrium strategy for the selfish peers. In evolutionary game, Evolutionary Stable Strategy (ESS) is a solution to game theoretic problems which is equivalent to the Nash equilibrium, but can be applied to the evolution of individuals behavior, which is defined as Definition 1. If all individuals are using a certain strategy, then it does not pay some to change to a different strategy if their expected payoff will be worse than the rest of the population in the system.

Definition 1. *A strategy x^* is an ESS, if and only if, $\forall x \neq x^*$, x^* satisfies:*

(1)equilibrium condition: $F(x^*,x^*) \geq F(x,x^*)$, and

(2)stability condition: if $F(x,x^*) = F(x^*,x^*)$, then $F(x^*,x) > F(x,x)$, where $F(x_1,x_2)$ is the utility obtained by a player when he/she uses strategy x_1 and the other player uses strategy x_2.

Since all peers are selfish and rational, they will cheat if their payoffs can be increased by cheating, which means that all peers are uncertain to the behavior and payoffs of the other peers. In such a case, peers will learn from strategy interaction in each round of game and try different strategies to improve their own utility. Hence, the percentage of peers who use a certain pure strategy may change during this process. Such kind of population evolution process can be modeled by replicator dynamics method, which is the basic dynamic mechanism of evolutionary game.

In a dynamic evolutionary game, an individual from a population (i.e., a player in the game), who is able to reproduce itself through the process of mutation and selection, is called a replicator. In this case, a replicator with a higher payoff can reproduce itself faster. When the reproduction process takes place over time, this can be modeled by using a set of ordinary differential equations called replicator dynamics. This replicator dynamics is important for an evolutionary game since it can provide information about the population, given a particular point in time, which can accurately describe the dynamic relationship between individual behavior payoff and group system evolution. Replicator dynamics refers to the growth rate of the proportion of the peers using certain pure strategy is equal to the difference between the payoffs obtained by using the strategy and the average payoffs of the peers within the group.

Assume that peers are in the homogeneous groups where the cost of all peers serving as a SGP is assumed to be the same. In replicator dynamics, it is assumed that a peer chooses pure strategy i from a finite set of strategies $B = \{SGP, OP\}$ including acting as a SGP or acting as an OP. Let n_i denote the number of individuals choosing strategy i, and let the total population size be $N = \sum_{i=1}^{2} n_i$. The proportion of individuals choosing strategy i is $x_i = n_i/N$, and it is referred to as the population share. The population state can be denoted by the vector $X = \{x_1, x_2\}$. The replicator dynamics can be described through the following differential equation:

$$\frac{dx_i}{dt} = x_i(t)\left[\overline{\pi}_i(t) - \overline{\pi}(t)\right], \tag{3}$$

where $\pi_i(t)$ is the average payoff obtained by the peers using pure strategy i, and $\overline{\pi}(t)$ is the average payoff of all peers within the group.

It can be seen from (3) that if using pure strategy i can get a higher payoff than the average level, the probability of a peer using strategy i will grow, and the growth rate is in proportion to the difference between the average payoff of using pure strategy i and the average payoff of all peers.

4. Analysis of the SGP Selection Evolutionary Game

According to (1) and (2), the average payoff of a peer if he/she choose to be a SGP can be computed by

$$\overline{\pi}_{SGP}(x)=\sum_{i=0}^{N-1}\binom{N-1}{i}x^i(1-x)^{N-1-i}\left[Pr(d_{i+1}\geq r)G-C_i\right] \tag{4}$$

where x is the probability of a peer being a SGP, and $\binom{N-1}{i}x^i(1-x)^{N-1-i}$ is the probability that there are i SGPs out of other $N-1$ peers. Similarly, the average payoff of a peer if he/she chooses to be an OP is given by (5):

$$\overline{\pi}_{OP}(x)=\sum_{i=1}^{N-1}\binom{N-1}{i}x^i(1-x)^{N-1-i}Pr(d_i\geq r)G. \tag{5}$$

According to (4) and (5), the average payoff of a peer is

$$\overline{\pi}(x)=x\overline{\pi}_{SGP}(x)+(1-x)\overline{\pi}_{OP}(x). \tag{6}$$

Substituting (6) back to replicator dynamics differential equation (3), for $i=SGP$, we have

$$\frac{dx}{dt}=x(1-x)[\overline{\pi}_{SGP}(x)-\overline{\pi}_{OP}(x)]. \tag{7}$$

At equilibrium point x^*, no player will deviate from the optimal strategy, which means $\frac{dx}{dt}|_{x^*}=0$, and we can get $x^*=0, x^*=1$, or x^* is the solutions to $\overline{\pi}_{SGP}(x)-\overline{\pi}_{OP}(x)=0$. However, since $\frac{dx}{dt}|_{x^*}=0$ is only the necessary condition for x^* to be ESS, we examine the sufficient condition for each ESS candidate and draw the following conclusions with the proofs shown in Theorem 1-3.

(1)$x^*=0$ is an ESS only when $Pr(d_1\geq r)G-C\leq 0$.
(2)$x^*=1$ is an ESS only when $Pr(d_N\geq r)G-Pr(d_{N-1}\geq r)G\geq C$.
(3)Let x^* be the solution to $\overline{\pi}_{SGP}(x)=\overline{\pi}_{OP}(x)$, and $x^*\in(0,1)$, then x^* is an ESS.

Lemma 1. *Let $g(x)=\overline{\pi}_{SGP}(x)-\overline{\pi}_{OP}(x)$, then $g'(x)<0$, $\forall x\in[0,1]$.*

Proof. According to (4) and (5), we have

$$g(x)=\sum_{i=0}^{N-1}\binom{N-1}{i}x^i(1-x)^{N-1-i}v_i-C, \tag{8}$$

where $v_i=[Pr(d_{i+1}\geq r)-Pr(d_i\geq r)]G$ is the additional gain by introducing one more SGP into the i-SGPs P2P streaming system.

For $\forall x\in[0,1]$, by taking the derivative of $g(x)$ over x, we have

$$\begin{aligned}g'(x)&=\sum_{i=0}^{N-1}\binom{N-1}{i}x^i(1-x)^{N-1-i}v_i-C\\&=\sum_{i=0}^{N-1}\binom{N-1}{i}[ix^{i-1}(1-x)^{N-1-i}\\&\quad-x^i(N-1-i)(1-x)^{N-2-i}]v_i\\&=\sum_{i=0}^{N-1}\binom{N-1}{i}[x^{i-1}(1-x)^{N-2-i}(i(1-x)\\&\quad-x(N-1-i))]v_i\\&=\sum_{i=0}^{N-1}\binom{N-1}{i}x^{i-1}(1-x)^{N-2-i}[i-x(N-1)]v_i.\end{aligned} \tag{9}$$

By introducing an integer i_1 that satisfies $i_1\leq(N-1)x<i_1+1$, we can get that

$$\begin{aligned}g'(x)&=\sum_{i=0}^{i_1}\binom{N-1}{i}x^{i-1}(1-x)^{N-2-i}[i-x(N-1)]v_i\\&+\sum_{i=i_1+1}^{N-1}\binom{N-1}{i}x^{i-1}(1-x)^{N-2-i}[i-x(N-1)]v_i.\end{aligned} \tag{10}$$

Since v_i is a decreasing function in terms of i, which means that $v_i\geq v_{i_1}$ when $\forall i\leq i_1$, and $v_i<v_{i_1}$ when $\forall i>i_1$. Therefore, according to (10), we have

$$\begin{aligned}g'(x)<&\sum_{i=0}^{i_1}\binom{N-1}{i}x^{i-1}(1-x)^{N-2-i}[i-x(N-1)]v_{i_1}\\&+\sum_{i=i_1+1}^{N-1}\binom{N-1}{i}x^{i-1}(1-x)^{N-2-i}[i-x(N-1)]v_{i_1}\\=&v_{i_1}\sum_{i=0}^{N-1}\binom{N-1}{i}x^{i-1}(1-x)^{N-2-i}[i-x(N-1)]\\=&v_{i_1}\frac{d\left[\sum_{i=0}^{N-1}\binom{N-1}{i}x^i(1-x)^{N-1-i}\right]}{dx}\\=&0\end{aligned}$$

and therefore, $g'(x)<0$, $\forall x\in[0,1]$.

Theorem 1. *The condition for $x^*=0$ to be an ESS is $Pr(d_1\geq r)G-C\leq 0$.*

Proof. According to (4-6), the average payoff that a peer using mixed strategy x and the other peers use mixed strategy $x^*=0$ can be written as:

$$\overline{\pi}(x,0)=\overline{\pi}_{OP}(0)+x[\overline{\pi}_{SGP}(0)-\overline{\pi}_{OP}(0)],$$

where $\overline{\pi}_{SGP}(0)=Pr(d_1\geq r)G-C, \overline{\pi}_{OP}(0)=0.$

(1)If $Pr(d_1 \geq r)G - C > 0$, i.e. $\overline{\pi}_{SGP}(0) > \overline{\pi}_{OP}(0)$, every peer will deviate to $x = 1$ to get $\overline{\pi}_{SGP}(0)$ rather than $\overline{\pi}_{OP}(0)$.
(2)If $Pr(d_1 \geq r)G - C < 0$, i.e. $\overline{\pi}_{SGP}(0) < \overline{\pi}_{OP}(0)$, every peer will select to stay at $x = 0$ to obtain $\overline{\pi}_{OP}(0)$ rather than $\overline{\pi}_{SGP}(0)$.
(3)If $Pr(d_1 \geq r)G - C = 0$, i.e. $\overline{\pi}_{SGP}(0) = \overline{\pi}_{OP}(0)$, then $g(0) = \overline{\pi}_{SGP}(0) - \overline{\pi}_{OP}(0) = 0$. According to Lemma 1, we know that $g'(x) < 0$, $\forall x \in [0,1]$, so $g(x) = \overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x) < g(0) = 0$, $\forall x \in [0,1]$. In such a case, we have $\overline{\pi}(0,x) = \overline{\pi}_{OP}(x) > \overline{\pi}(x,x) = \overline{\pi}_{OP}(x) + x(\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x))$, which means $x^* = 0$ is an ESS according to Definition 1.

Therefore, $x^* = 0$ is an ESS only when $Pr(d_1 \geq r)G - C \leq 0$.

Theorem 2. *The condition for $x^* = 1$ to be an ESS is $Pr(d_N \geq r)G - Pr(d_{N-1} \geq r)G \geq C$.*

Proof. According to (4-6), the average payoff that a peer use mixed strategy x and the other peers use mixed strategy $x^* = 1$ can be written as:

$$\overline{\pi}(x,1) = \overline{\pi}_{OP}(1) + x[\overline{\pi}_{SGP}(1) - \overline{\pi}_{OP}(1)],$$

where $\overline{\pi}_{SGP}(1) = Pr(d_N \geq r)G - C$, $\overline{\pi}_{OP}(1) = Pr(d_{N-1} \geq r)G$.

(1)If $Pr(d_N \geq r)G - Pr(d_{N-1} \geq r)G < C$, i.e. $\overline{\pi}_{OP}(1) > \overline{\pi}_{SGP}(1)$, every peer will deviate to $x = 0$ to get $\overline{\pi}_{OP}(1)$ rather than $\overline{\pi}_{SGP}(1)$.
(2)If $Pr(d_N \geq r)G - Pr(d_{N-1} \geq r)G > C$, i.e. $\overline{\pi}_{OP}(1) < \overline{\pi}_{SGP}(1)$, every peer will select to stay at $x = 1$ to obtain $\overline{\pi}_{SGP}(1)$ rather than $\overline{\pi}_{OP}(1)$.
(3)If $Pr(d_N \geq r)G - Pr(d_{N-1} \geq r)G = C$, i.e. $\overline{\pi}_{OP}(1) = \overline{\pi}_{SGP}(1)$, then $g(1) = \overline{\pi}_{SGP}(1) - \overline{\pi}_{OP}(1) = 0$. According to Lemma 1, we know that $g'(x) < 0$, $\forall x \in [0,1]$. In such a case, we have $\overline{\pi}(1,x) = \overline{\pi}_{OP}(x) + 1 \times (\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x)) > \overline{\pi}(x,x) = \overline{\pi}_{OP}(x) + x(\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x))$, which means $x^* = 1$ is an ESS according to Definition 1.

Therefore, $x^* = 1$ is an ESS only when $Pr(d_N \geq r)G - Pr(d_{N-1} \geq r)G \geq C$.

Theorem 3. *If $x^* \in (0,1)$ is a solution to equation $\overline{\pi}_{SGP}(x) = \overline{\pi}_{OP}(x)$, then x^* is an ESS.*

Proof. Let $\overline{\pi}_i(x,x^*)$ be the average payoff of player i that uses mixed strategy x and the other peers use mixed strategy x^*. Then we have

$$\overline{\pi}_i(x,x^*) = x\overline{\pi}_{SGP}(x^*) + (1-x)\overline{\pi}_{OP}(x^*). \qquad (11)$$

Since x^* is a solution to $\overline{\pi}_{SGP}(x) = \overline{\pi}_{OP}(x)$, we have $\overline{\pi}_{SGP}(x^*) = \overline{\pi}_{OP}(x^*)$, and (11) becomes

$$\overline{\pi}_i(x,x^*) = \overline{\pi}_{SGP}(x^*) = \overline{\pi}_i(x^*,x^*), \qquad (12)$$

which means x^* satisfies the equilibrium condition defined in Definition 1.

Moreover, according to (6), we have $\overline{\pi}_i(x,x) = \overline{\pi}_{OP}(x) + x(\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x))$ and $\overline{\pi}_i(x^*,x) = \overline{\pi}_{OP}(x) + x^*(\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x))$, therefore, we have $\overline{\pi}_i(x^*,x) - \overline{\pi}_i(x,x) = (x^* - x)(\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x))$.

From Lemma 1, we know that $g(x) = \overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x)$ is monotonically decreasing. Since $\overline{\pi}_{SGP}(x^*) = \overline{\pi}_{OP}(x^*)$, so when $x < x^*$, we have $\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x) > 0$; and when $x > x^*$, we have $\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x) < 0$. Therefore, for $\forall x \neq x^*$, we have $(x^* - x)(\overline{\pi}_{SGP}(x) - \overline{\pi}_{OP}(x)) > 0$, i.e. $\overline{\pi}_i(x^*,x) > \overline{\pi}_i(x,x)$, $\forall x \neq x^*$, which means x^* satisfies the stability condition defined in Definition 1. So we know that x^* is an ESS.

5. A Distributed Q-Learning Algorithm for ESS

From the previous section, we can see that the ESS can be found by solving the replicator dynamics equations. However, solving the replicator dynamics equations require the exchange of strategies adopted by other peers and their private information. In such a case, a Q-learning approach [13] which is a type of reinforcement learning (i.e., learning by interaction) is used. With this ability to learn, complete payoff information of other users in the same group is no longer required for SGP selection. In this section, we will present a distributed Q-learning algorithm that can gradually converge to ESS without information exchange.

We first discretize the replicator dynamics equation as

$$x_i(t+1) = x_i(t) + [\overline{\pi}_i(t) - \overline{\pi}(t)]x_i(t), \qquad (13)$$

where t is the slot index and $x_i(t)$ is the probability of peer i being a SGP during slot t. Here, we assume that each slot can be further divided into M subslots and each peer can choose to be a SGP or an OP at the beginning of each subslot.

From (13), we can see that in order to update $x_i(t+1)$, we need to first compute $\overline{\pi}_i(B, x_{-i}(t))$ and $\overline{\pi}_i(x_i(t))$. Let us define an indicator function $o_i(t,k)$ whose value equal to 1 if player i choose to be a SGP at the beginning of kth subslot in time slot t. Otherwise, the function value is 0.

The directly payoff of player i at subslot k in slot t can be computed by (14):

$$\pi_i(t,k) = \begin{cases} G - C, & \text{if player } i \text{ chooses to be a SGP, and } r^t \geq r, \\ -C, & \text{if player } i \text{ chooses to be a SGP, and } r^t < r, \\ G, & \text{if player } i \text{ chooses to be an OP, and } r^t \geq r, \\ 0, & \text{if player } i \text{ chooses to be an OP, and } r^t < r. \end{cases} \qquad (14)$$

where r^t is the total download rate of the SGPs and r is the video source rate.

Then, $\overline{\pi}_i(B, x_{-i}(t))$ can be approximated using (15):

$$\overline{\pi}_i(B, x_{-i}(t)) = \frac{\sum_{k=1}^{M} \pi_i(t,k) o_i(t,k)}{\sum_{k=1}^{M} o_i(t,k)}, \tag{15}$$

Similarly $\overline{\pi}_i(x_i(t))$ can be approximated as

$$\overline{\pi}_i(x_i(t)) = \frac{1}{M} \sum_{k=1}^{M} \pi_i(t,k). \tag{16}$$

In the distributed Q-learning algorithm, Q-value (i.e., $Q(t)$) is used to maintain the knowledge about each peer, and the decision can be made based on this knowledge. The SGP-selection algorithm can be described as follows.

Algorithm 1: A Distributed Q-learning SGP Selection Algorithm for ESS	
Step 1:	Given the slot index $t = 0$, each peer initializes $x_i(t) = x_i(0)$ with a random between [0,1].
Step 2:	Initialize Q-value associated with $Q(t) = 0$ for all peers in super groups.

Step 3:	**loop**
	During slot t, for $k = [1 : M]$
	if $x_i(t) \leq \gamma$ **then**
	player i randomly choose to be a SGP to serve as a super peer to download streaming chunks from the peers outside the group with download rate $r_i(t,k)$ or to be an OP to download streaming chunks from the SGP.
	else
	Choose strategy $b^* = arg\max_b Q(k)$
	end if
	player i computes the indicator function $o_i(t,k)$ and his/her directly payoff $\pi_i(t,k)$ using(14).

Step 4:	player i computes average payoff $\overline{\pi}_i(B, x_{-i}(t))$ and $\overline{\pi}_i(x_i(t))$.
Step 5:	Updates the probability of being a SGP $x_i(t+1)$ using (13).
Step 6:	Update $Q(k+1) = (1-\lambda)Q(k)$ $+\lambda\left(\overline{\pi}_i(x_i(t)) + \beta \max_b Q(k)\right)$
Step 7:	endloop for all peers in super group.

In this SGP-selection algorithm, a peer performs the exploration step with probability γ, and λ denotes the learning rate that is used to control the speed of adjustment of the Q-value. A new Q-value $Q(k+1)$, which is the expected payoff for the future iterations, is obtained based on the previous value $Q(k)$ along with the new observed payoff $\overline{\pi}_i(x_i(t))$. Here, the new observed payoff is biased by the outcome of choosing the best action based on the available knowledge (i.e., $\max_b Q(k)$).

In each hybrid hierarchical P2P live streaming super group, the complexity of updating the probability of being a SGP among N peer nodes is $O(N)$. The complexity and computational time increase when the number of peers in the same super group increases. However, the SGP selection algorithm based on ESS can be performed offline to update the peer nodes strategies, which means that no real-time update is needed in real implementation. The values of indicator function and immediate utility can be sent from peer nodes which are collected periodically by the central coordinator at the channel server, and then the action taken by peer nodes is analyzed and updated according to Algorithm 1.

6. Experimental Results

To the best of our knowledge, there does not exist appropriate algorithms in the literature that can be directly used to solve the problem under study. For evaluation purposes, we consider the traditional randomly select super peers method, which is denoted as Random-SGP for performance comparison. In Random-SGP, each peer acts as an individual and randomly selects some peers from its partner list for downloading video chunks. Such a protocol has been widely used in the existing P2P systems, such as PPLive [15] and Cool Streaming[16].For convenience, in the rest of this paper, we denote the proposed ESS-based approach as ESS-SGP.

In this section, we describe the experiment results which are conducted on the LStream[14] P2P live streaming platform deployed in the campus of Zhengzhou University. In our experiments, there are about 1000 users in the hybrid layered P2P live streaming system. The video is initially stored at an original video server with upload bandwidth $r = 1Mbps$. The request round is 1 second and the relay buffer length is 30 seconds and the streaming fragment size is $L = 1KB$.

In the first experiment, assume that there are 20 homogeneous peers in one super group, their gain is $G = 1$ and their cost is $C = 0.1$. The first experiment compares the differences between ESS-SGP method and traditional Random-SGP method in the aspect of the obtained social welfare, which is defined by the sum of the utility obtained by all peers. As it is shown in Fig. 2, ESS-SGP method obtains a stable Nash equilibrium and chooses suitable numbers of SGPs through the collaboration among the peers in the same super group. It maintains a relatively higher social welfare level under the condition of relatively higher video source rate r. But the social welfare level of traditional Random-SGP method decreases linearly and rapidly with the source rate increasing.

In the second experiment, we evaluate the convergence property of the ESS-SGP. In Fig. 3, we show that the replicator dynamic of the SGP selection game with homogeneous peers with $r = 1Mbps$. We can see that starting from high initial values, all peers gradually

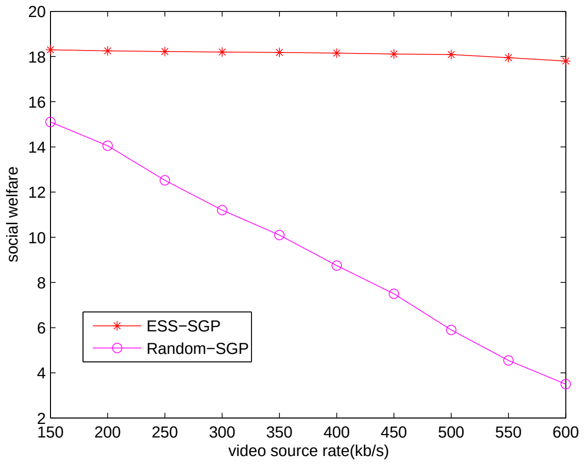

Figure 2 The social welfare comparison between ESS-SGP and Random-SGP.

reduce their probabilities of being a SGP since being an OP more often can bring a higher payoff. However, because too low a probability of being a SGP may increases the chance of having no peer be a SGP, the probability of being a SGP will finally converge to a certain value which is determined by the number of peers.

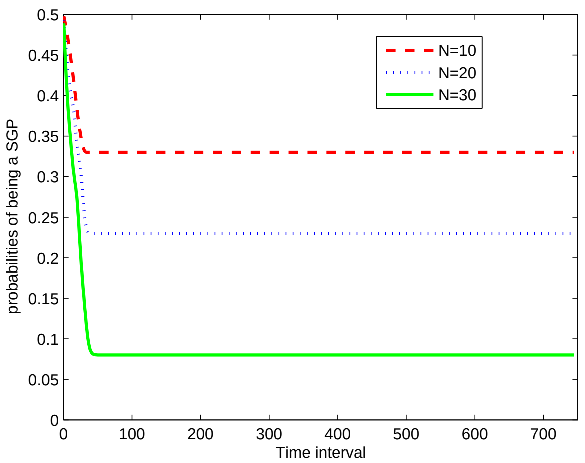

Figure 3 Behavior dynamics of a homogeneous group of peers.

The third experiment compares the upload capacity of the P2P network using ESS-SGP algorithm and traditional Random-SGP method. As it is shown in Fig. 4, when the nodes number up to about 2100, the load of the SGP nodes close to saturation, and when the system scale continues to increase, the increase in the upload capacity of the system is not obvious. On the contrary, ESS-SGP algorithm eliminates the system bottlenecks and the upload capacity of the P2P live streaming system increased steadily with the number of peers increasing.

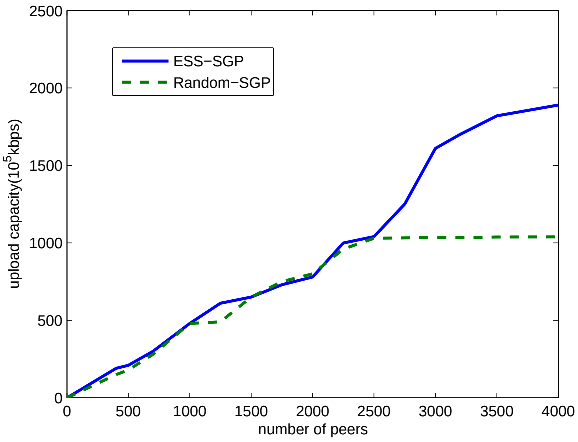

Figure 4 The upload capacity of P2P streaming system comparison between ESS-SGP and Random-SGP.

7. Conclusion

In this paper, we propose a super group peer selection game model based on evolutionary game, which is used in hierarchical hybrid P2P streaming system to address the network inefficiency problem encountered by the traditional SGP random selected scheme. By deriving the ESS for every peer, we further propose a distributed Q-learning algorithm for each peer to converge to the ESS by learning from his/her own past payoff history. The experiment results indicate that compared with the traditional Random-SGP scheme, the proposed algorithm achieves much better social welfare, and enable the upload capacity of the whole P2P system to increase stably with the peer number increasing. In further study, we will take into account the heterogeneity of the peers within peer groups, which are much closer to the actual situation of P2P streaming system.

Acknowledgement

The first author acknowledges the financial support by the National major science and technology program of Ministry of Industry and Information Technology of China under Grant No.2009ZX03001 -018-01 with Zhengzhou University. We would like to thank the project team in the Henan Provincial Key Laboratory on Information Network for their help with the implementation, deployment, and experiments of the live media streaming system in real networks.

The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] V. Lo, D. Zhou, Y. Liu, C. GauthierDickey, and J. Li, Scalable supernode selection in peer-to-peer overlay networks, Proceedings of Second International Workshop on Hot Topics in Peer-to-Peer Systems 2005(HOT-P2P 2005), San Diego, CA, 18-25 (2005).

[2] C. Yu, X. Gou, C. Zhang and Y. Ji, Study on Supernode Election Algorithm in P2P Network Based upon District Partitioning, Proceedings of International Conference on Communication Software and Networks 2009 (ICCSN '09), Macau, China, 196-199 (2009).

[3] J. Kratica, J. Kojić, D. Tošić, V. Filipović, and D. Dugošija, Two Hybrid Genetic Algorithms for Solving the Super-Peer Selection Problem, Advances in Intelligent and Soft Computing **58**, 337-346 (2009).

[4] S.H. Min and D.S. Cho, Super-Peer selection based framework using dynamic capacity and similarity, Computer and Information Sciences(ISCIS 2006), Springer-Verlag, Berlin, Heidelber **4**, 803-812 (2006).

[5] L. Gao, and M. Peng, Optimal Superpeer Selection Based on Load Balance for P2P File-sharing System, Proceedings of International Joint Conference on Artificial Intelligence 2009(JCAI '09), Haikou, China, 92-95 (2009).

[6] S.H. Min, J.Holliday, and D.S. Cho, Optimal Super-peer Selection for Large-scale P2P System, Proceedings of International Conference on Hybrid Information Technology 2006 (ICHIT '06), Jeju Island, Korea, 588-593 (2006).

[7] Y. Jin, Y. Liu, and H. Zhao, Trust-based supernode selection in peer-to-peer systems, Proceedings of 2nd International Conference on Future Computer and Communication (ICFCC 2010), Wuhan China, 285-289 (2010).

[8] J. Halpern, Computer science and game theory: A brief survey, New Palgrave Dictionary of Economics. Palgrave Macmillan, 2nd edition, 24-26 (2008).

[9] S. Roy, C. Ellis, S. Shiva, A Survey of Game Theory as Applied to Network Security, Proceedings of 43rd Hawaii International Conference on System Sciences, Poipu, Hawaii, 1-10 (2010).

[10] XK Lu, QS Wu, RZ Li and YY Lin, On Tree Construction of Super Peers for Hybrid P2P Live Media Streaming, Proceedings of 19th International Conference on Computer Communications and Networks (ICCCN 2010), ETH Zurich, Switzerland, 1-6 (2010).

[11] Y. Huang, Tom Z.J.Fu, D-M Chiu, John C.S.Lui, C.Huang, Challenges, design and analysis of a large-scale p2p-vod system, SIGCOMM Computer Communication Review **38**, NY, USA, 375-388 (2008).

[12] J. M. Smith, Evolutionary and the theory of games, Cambridege University Press, 1982.

[13] D. Fudenberg and D. K. Levine, The Theory of Learning in Games, Cambridge, MA: MIT Press, 1998.

[14] LStream P2P Live Streaming System, http://lstream.ha.edu.cn/.

[15] S. Spoto, R. Gaeta, M. Grangetto, M.Sereno, Analysis of PPLive through active and passive measurements, Proceedings of IEEE International Symposium on Parallel and Distributed Processing 2009 (IPDPS 2009), 1-7 (2009).

[16] B. Li, S. Xie, Y. Qu, G.Y.Keung, C. Lin, JC.Liu, XY.Zhang, Inside the new coolstreaming: Principles, measurements and performance implications, Proceedings of 27th IEEE Conference on Computer Communications (INFOCOM 2008), Phoenix, AZ USA, 1031-1039 (2008).

Jing Chen received the B.S. and M.S. degree in computer science from Zhengzhou University in China in 1999 and 2002 respectively, and she is pursuing her Ph.D. degree in computer science from Xi'an Jiaotong University. Her research interests include peer -to-peer streaming, wireless multimedia, game theory, and network security.

Yangjie Cao received his M.S. degree in Computer Science from Zhengzhou University, China, in 2003, and his Ph.D. degree in Computer Science from Xi'an Jiaotong University, China, in 2012. Currently he is a lecturer in Software Technology School, Zheng-zhou University, China. His research interests mainly include adaptive scheduling and high-performance computing.

Li Li received the B.S. and M.S. degrees in computer science from Shanxi Normal University, Xian, China, in 2002 and 2005 respectively, and the Ph.D. degree in systems engineering from Xian Jiaotong University, Xian, China, in 2012. Her research interests include P2P streaming system, network security, reconfigurable network and survivable network.

Xiaoshe Dong was born in Xi'an, Shanxi Province, China in 1963. He is a member of IEEE, and the professor and also a Ph.D. supervisor now at School of Electronic and Information Engineering in Xi'an Jiaotong University, China. His research interests include high performance computing, trusted computing and wireless embedded multimedia system.