

分层 P2P 实时流媒体系统中基于演进博弈的超级组对等节点选择算法

陈静^{1,2}, 周兵², 任晨珊², 董小社¹

(1. 西安交通大学 电子与信息工程学院, 陕西 西安 710049; 2. 郑州大学 信息工程学院, 河南 郑州 450052)

摘要: 针对分层混合式 P2P 实时流媒体系统中的超级组对等节点选择问题, 提出了一种基于演进博弈论的分布式超级组对等节点选择算法。首先使用演进博弈框架对 P2P 实时流媒体系统中的超级组对等节点选择问题进行了建模, 并对其演进稳定策略进行了求解分析。然后通过得到的混合策略设计了一种基于 Q-Learning 的分布式 SGP 选择算法 ESS-SGP, 以达到激励对等组内的对等节点之间互相协作的目的。实验结果表明, 与传统的随机选择超级组对等节点方法相比, 使用 ESS-SGP 算法所选择产生的超级组对等节点能够稳定地为系统提供更高的流媒体上传速率, 使整个 P2P 流媒体系统的上传能力随着对等节点数目的增加而稳定增加, 并为每个对等节点带来更好的收益。

关键词: P2P 流媒体; 超级组对等节点选择; 演进博弈; 演进稳定策略

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2013)12-0094-10

Super group peer selection algorithm in hybrid layered P2P streaming system based on evolutionary game

CHEN Jing^{1,2}, ZHOU Bing², REN Chen-shan², DONG Xiao-she¹

(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)

Abstract: Due to high efficiency and good scalability, hybrid layered P2P architecture receives more and more importance in P2P research and P2P streaming application fields recently. The problems about super peers selection are highly challenging because super peers must be selected from huge and dynamically changing network. A distributed super group peer(SGP) selection algorithm based on evolutionary game is proposed in this paper, in which the SGP selection procedure is modeled based on evolutionary game framework firstly and its evolutionarily stable strategies are analyzed. Moreover, a distributed Q-Learning algorithm (ESS-SGP) according to the mixed strategies by analysis is proposed for the peers to converge to the ESSs based on their own payoff history. Compared to the traditional random selected SGP P2P schemes, experiments results show that the proposed ESS-SGP algorithm achieves much better performance in terms of social welfare, average upload rate of SGPs, and keeps the upload capacity of the P2P streaming system increasing steadily with the number of peers increasing.

Key words: P2P streaming system; super group peer selection; evolutionary game; evolutionarily stable strategy

1 引言

根据覆盖网络结构类型的不同, 典型的 P2P 视频流媒体系统分为基于树状(tree-push)拓扑、基于网状(mesh-pull)拓扑和基于分层混合拓扑 3 种模型^[1]。其中, 分层混合拓扑模型根据对等节点的性能差

异, 将 P2P 网络中的对等节点分为超级节点和普通节点。该模型综合了树状拓扑和网状拓扑结构的优点, 让能力强的节点充分参与到系统中, 具有更高的可扩展性和稳定性, 成为最近研究的重点。但是, 在分层混合式 P2P 流媒体系统中, 如何从众多动态、匿名、自治的异构普通节点中选取部分节点作为

收稿日期: 2012-12-13; 修回日期: 2013-05-25

基金项目: 国家自然科学基金资助项目 (61379079)

Foundation Item: The National Natural Science Foundation of China (61379079)

超级节点为其他节点提供服务是一个非常关键的问题^[2]，关系到整个 P2P 流媒体系统的效率和用户的体验质量(QoE, quality of experience)。

目前，在分层混合 P2P 流媒体系统中的超级对等节点选择方面有了许多研究，但是大多数研究只关注到了节点的能力^[2-5]，如内存空间、CPU 周期或网络带宽等，很少研究工作关注于对等节点的行为属性。超级对等节点要为组内的其他普通节点提供资源查询和路由服务，并为他们上传视频流媒体分片数据，节点的计算和存储空间消耗很大，有可能造成节点性能的明显下降。文献[2]中提出的用于非结构化 P2P 网络中选取超级对等节点的 H₂O (hierarchical 2-level overlay)协议在超级对等节点的选择过程中考虑到了信任、安全路径和路由性能等方面的因素，但是超级对等节点间的通告报文通过洪泛方式分发，造成了很大的网络开销。Jin Y 等人在文献[6]中提出了一种基于代理信任的超级对等节点选择方法，在一定程度上降低了文献[2]中的网络洪泛方面的开销，但是这种方法需要先验知识才能做出是否信任某个节点的决策。Yu C 等人在文献[7]中提出了一种基于对等节点 IP 地址进行区域划分的超级对等节点选择方法，将整个 P2P 覆盖网络划分成 k 个小的区域，并使用分布式并行算法分别在各个区域内进行超级节点的选举，以降低通信和计算的复杂度。Gao Z W 等人在文献[8]中提出了一种基于自信息理论的超级节点选择算法，该选择算法根据每个对等节点上维护的信息向量进行加权平均，从而选出节点能力和节点在线时间最好的节点作为超级对等节点。Wolf S 等人将 P2P 网络中的超级对等节点选择问题看作是一个 Hub 选址优化问题，在文献[9]中证明了该问题是一个 NP-hard 问题，并在文献[10]中提出了一种基于进化算法和局部搜索的解决方法。Kratika J 等人针对超级对等节点选择问题在文献[11]中提出了 2 种混合遗传算法，以减少整个 P2P 网络中的通信开销。该算法对 P2P 网络中的每一个对等节点进行了两段式基因编码并为之设计适合的目标函数和修正后的遗传算子，并采用了基于局部搜索方式的启发算法以减少对等节点间的通信开销。但上述工作均是针对一般的 P2P 网络进行研究，并未考虑 P2P 实时流媒体网络的特点。Wang S L 等人在文献[12]中提出了一种基于层次分析法(AHP, ana-

lytic hierarchy process)的超级对等节点选择机制，该机制将对等节点的兴趣向量相似度、带宽以及网络传输往返时延 RTT 作为超级对等节点选择时的决策变量进行决策分析，但该机制没有考虑对等节点的行为属性。

近年来，博弈论作为一种对 2 个或多个决策者间的策略交互进行建模的数学工具，近年来在计算机科学的各个领域中也得到了越来越多的关注^[13-16]。博弈论假设参与博弈的每个参与者是“理性的”，即每个参与者理性地选择符合使自己利益最大化的策略。在文献[14]中，博弈论被用来解决移动 ad-hoc 网络中的入侵检测、DDoS 攻击问题以及因特网蠕虫攻击等网络安全问题。P2P 视频流媒体系统中的对等节点是理性而自私的，他们可以根据自己的需要和其他对等节点的行为来进行决策。Chisci L 等人在文献[15]中提出了一种基于演进博弈的视频流联合资源分配和路由算法，以解决 P2P 流媒体系统中的“公地悲剧”问题，但没有对提出的技术进行理论方面的分析。Bioglio V 等人在文献[16]中针对目前 P2P 平台中存在的重叠网络与底层网络拓扑失配的问题，提出了一种可以用于设计因特网服务提供商(ISP)策略的博弈理论框架，使得 ISP 间的流量达到最小化，以达到高效可靠地支持 P2P 流媒体服务的目的。该框架指出了均衡点的存在性，对这些均衡点进行精炼时可能采取的策略，以及从演进博弈的角度如何达到均衡。但是文献[16]中提出的算法在具体实现时需要 ISP-tracker 与 super-tracker 之间互相信任。本文在已有的工作基础上，提出了一种基于演进博弈的分布式超级组对等节点选择算法。首先使用演进博弈框架对 P2P 实时流媒体系统中的超级对等节点选择问题进行了建模，并对其演进稳定策略 ESS (evolutionarily stable strategy) 进行了求解分析。然后通过得到的混合策略设计了一种基于 Q-learning 的分布式超级组对等节点选择算法，达到激励对等组内的对等节点之间互相协作的目的。最后在郑州大学校园中部署的 P2P 实时流媒体平台 LStream^[17]上进行了算法比较的实验结果，实验结果表明，与传统的随机选择 SGP 方法相比，ESS-SGP 能够稳定地为系统提供更高的流媒体上传速率，能够使整个 P2P 系统的上传能力随着对等节点数目的增加而稳定增加，并能够为每个对等节点带来更好的收益。

2 系统模型和效用函数

2.1 系统模型

如图 1 所示, 在分层混合式 P2P 流媒体系统的边缘有多个对等节点希望同时观看一段实时的视频流, 希望某个频道节目的流媒体分片数据尽量在一个对等节点组内共享, 以减少不必要的网络流量^[18]。为此, 系统将它们按照相对的地理位置划分为多个对等节点组(peer group)。在一个组内部, 每个对等节点可以选择做一个超级组对等节点(SGP, super group peer)或者做一个普通对等节点(OP, ordinary peer)。如果选择作为超级组对等节点 SGP, 那么该对等节点不但需要作为一个客户从其他组的超级组对等节点那里下载视频流媒体分片数据, 而且还需要作为一个服务器为自己组内的普通对等节点和其他组的超级组对等节点上传视频流媒体分片数据。反之, 它只需要在同一个组内为其他对等节点上传或从 SGP 那里下载需要的流媒体分片。假定组内的对等节点之间的上传和下载带宽远大于对等组之间的上传和下载带宽, 在这种情况下, 对等节点出于自私的本性会倾向于选择做一个普通对等节点。然而, 从另外一方面来看, 如果没有足够的对等节点选择做超级组对等节点就会存在实时流媒体播放质量降低的风险, 因为有可能出现没有足够的超级组对等节点从其他对等组下载视频流媒体分片数据的情况。本文对对等组的对等节点建立流媒体分发演进博弈模型, 希望激励对等节点组内的对等节点之间互相协作, 选择作为 SGP 节点来互相提供服务, 以获取更好的流媒体性能。

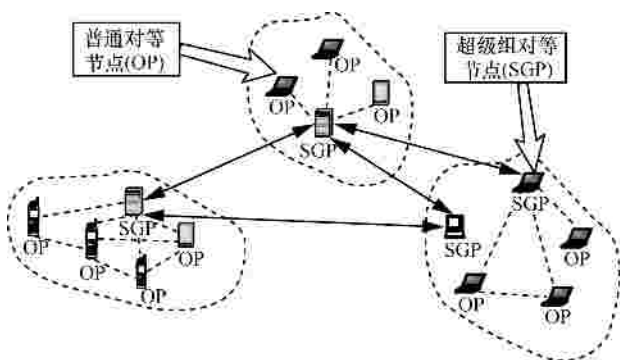


图 1 分层混合式 P2P 实时流媒体系统边缘中的对等节点组

2.2 效用函数

在分层混合式移动 P2P 流媒体系统中, 超级对

等节点既作为客户从其他对等节点那里下载所需要的视频流媒体分片数据, 还要作为服务器为其他对等节点上传视频流媒体分片数据。当一个对等节点从其他对等节点那里通过下载视频流媒体分片数据获益的同时, 它也会因为需要为其他对等节点上传视频流媒体分片数据而造成一定的开销, 这个开销可能是在上传视频流媒体分片时消耗的资源, 如带宽、内存缓冲区等。

假定对等组内有 N 个对等节点 u_1, u_2, \dots, u_N , 其中, 有 s 个对等节点愿意作为超级组对等节点 SGP 从组外的对等节点那里下载流媒体分片数据。设超级组对等节点 SGP 与相应的组外的对等节点之间的下载速率为两者之间的传输速率, 并将 s 个 SGP 的下载速率记为 r_1, r_2, \dots, r_s , 则这个组对等节点的总下载速率为

$$d_s = \sum_{i=1}^s r_i \tag{1}$$

因为这 s 个超级组对等节点 SGP 是随机、独立地选择组外的对等节点来进行流媒体分片数据的下载, 因此, 下载速率 r_i 是随机变量。根据文献 [19], 对等节点下载带宽的累积分布函数 CDF 可以建模成一个线性函数, 对等节点的下载带宽的概率密度函数 PDF 可以被看作是一个均匀分布, 即 r_i 是均匀分布的。

为清楚起见, 先考虑对等节点上没有缓冲的简单情况。在对等节点上没有缓冲的情况下, 如果总的下载速率 d_s 不小于视频源的源速率 r , 则组内的对等节点可以获得实时流媒体的效果, 这时所有的对等节点可以获得一定的收益 G 。否则, 就会存在一些延迟, 在这种情况下假定组内对等节点的收益为 0。因此, 给定总下载速率 d_s 和视频源速率 r , 如果对等节点 u_i 选择作为一个超级组对等节点 SGP, 则它的效用函数为

$$U_{SGP,i}(s) = \Pr(d_s \geq r)G - C_i, \quad \forall s \in [1, N] \tag{2}$$

其中, C_i 是对等节点 u_i 作为一个超级组对等节点 SGP 时的开销, $\Pr(d_s \geq r)$ 是对等节点获得实时流媒体效果的概率, 其计算方法参见定理 1。

定理 1 如果 r_1, r_2, \dots, r_s 是在 $[r^L, r^U]$ 上独立均匀分布的随机变量, 则概率 $\Pr(d_s \geq r)$ 为

$$\Pr(d_s \geq r) = \frac{1}{2s!} \sum_{m=0}^s (-1)^m \binom{s}{m} (s-m)^s -$$

$$\text{sgn}(\hat{r} - m) (\hat{r} - m)^s \quad (3)$$

当 s 足够大时，概率 $\Pr(d_s > r)$ 可以近似计算为

$$\Pr(d_s > r) \approx Q\left(\frac{\hat{r} - \frac{s}{2}}{\sqrt{\frac{s}{12}}}\right) \quad (4)$$

其中， $\hat{r} = \frac{r - sr^L}{r^U - r^L}$ ， $Q(x)$ 是高斯尾函数 $\int_x^\infty \frac{1}{\sqrt{2\pi}} \cdot$

$\exp\left(-\frac{x^2}{2}\right) dx$ 。定理 1 的具体证明如下。

证明 令 $\hat{r}_i = \frac{r_i - r^L}{r^U - r^L}$ ，则 $\hat{r}_1, \hat{r}_2, \dots, \hat{r}_s$ 是在区间 $[0,1]$ 上独立同分布的均匀分布的随机变量。 \hat{r}_i 的特征函数为

$$F(t) = \frac{i(1 - e^{-it})}{t} \quad (5)$$

令 $\hat{y}_s = \sum_{i=1}^s \hat{r}_i$ ，则 \hat{y}_s 的特征函数可由式(16)进行计算

$$F_{\hat{y}_s}(t) = \left(\frac{i(1 - e^{-it})}{t}\right)^s \quad (6)$$

则 \hat{y}_s 的密度函数为

$$\begin{aligned} f_{\hat{y}_s}(y) &= F_t^{-1}\left[\left(\frac{i(1 - e^{-it})}{t}\right)^s\right](y) \\ &= \frac{1}{2(s-1)!} \sum_{l=0}^s (-1)^l \binom{s}{l} \text{sgn}(y-l)(y-l)^{s-1} \end{aligned} \quad (7)$$

因为 $\Pr(y_s > r) = \Pr(\hat{y}_s > \hat{r})$ ，根据式(7)，有

$$\begin{aligned} \Pr(y_s > r) &= \Pr(\hat{y}_s > \hat{r}) = \int_{\hat{r}}^\infty f_{\hat{y}_s}(y) dy \\ &= \frac{1}{2s!} \sum_{l=0}^s (-1)^l \binom{s}{l} [(s-l)^k - \text{sgn}(\hat{r}-l) (\hat{r}-l)^s] \end{aligned} \quad (8)$$

当 s 足够大时，根据中心极限定理， \hat{y}_s 可以被近似为高斯分布 $N\left(\frac{s}{2}, \frac{s}{12}\right)$ 。因此，有

$$\Pr(y_s > r) = \Pr(\hat{y}_s > \hat{r}) \approx Q\left(\frac{\hat{r} - \frac{s}{2}}{\sqrt{\frac{s}{12}}}\right) \quad (9)$$

由于对等组内的上传和下载带宽都很大，向同一个组内的其他对等节点上传流媒体分片数据的开销可以忽略。在这种情况下，如果一个对等节点不选择作为一个超级组对等节点 SGP，而是选择作为一个普通对等节点 OP，则该对等节点没有开销，其效用函数为

$$U_{op,i}(s) = \begin{cases} \Pr(d_s > r)G, & s \in [1, N-1] \\ 0, & s=0 \end{cases} \quad (10)$$

3 超级组对等节点选择演进博弈

为了给对等节点提供一个健壮的均衡策略，本文采用了演进稳定策略 ESS 的概念^[20,21]，其定义如下。

定义 1 一个策略 x^* 是一个演进稳定策略，当且仅当 $\forall x \neq x^*, x^*$ 满足：

1) 均衡条件， $U_i(x, x^*) = U_i(x^*, x^*)$ ；

2) 稳定条件，如果 $U_i(x, x^*) = U_i(x^*, x^*)$ ，则有 $U_i(x, x) < U_i(x^*, x)$ ，其中， $U_i(b_1, b_2)$ 是当参与者 i 使用策略 b_1 ，而另外一位参与者使用策略 b_2 时，参与者 i 所获得的效用。

在 P2P 实时流媒体系统中，所有的对等节点对于其他对等节点的行为和效用都是不确定的。在这种情况下，为了提高自己的效用，对等节点在每一轮博弈中互相学习交互策略，并在每一轮博弈中采取不同的策略。在这个过程中，使用某个特定的纯策略的对等节点的百分比数可能会有所改变。这样的一种群体演进过程可以通过复制者动态(replicator dynamics)方法来进行建模。复制者动态是演进博弈的基本动态机制，能够比较准确地描述个体行为收益与群体系统演进的动态关系，它指的是使用某种纯策略的对等节点的个数所占比例的增长率，与使用该策略时所得收益与组内对等节点平均收益之差成正比。

假定对等节点选择作为超级组对等节点 SGP 时的开销相同。令 x_b 表示某个对等节点使用纯策略 $b \in ?$ 的概率，其中， $? = \{SGP, OP\}$ 是一个纯策略集合，元素 SGP 表示对等节点采用“作为超级组对等节点 SGP”的策略，而元素 OP 则表示对等节点采用“作为普通对等节点 OP”的策略。则 x_b 的演进动态性可以通过下面的微分方程进行描述：

$$\frac{dx_b}{dt} = x_b [\bar{U}(x_b, x_{-b}) - \bar{U}(x_b)] \quad (11)$$

其中, x_b 表示某个对等节点使用纯策略 $b \in ?$ 的概率, $\bar{U}(x_b, x_{-b})$ 是使用纯策略 b 的对等节点所得的平均收益, x_{-b} 是没有使用纯策略 b 的对等节点的集合, $\bar{U}(x_b)$ 是组内所有对等节点的平均收益。

从式(11)可以看出, 如果采用纯策略 b 能够得到比平均水平更高的收益, 那么对等节点使用策略 b 的概率将会增加, 增长率与使用纯策略 b 的对等节点的平均收益与所有对等节点的平均收益之差成正比。

4 超级组对等节点选择演进博弈分析

根据式(2)和式(10), 当对等节点选择作为超级组对等节点 SGP 时的平均收益可以通过式(12)进行计算。

$$\bar{U}_{SGP}(x) = \sum_{i=0}^{N-1} \binom{N-1}{i} x^i (1-x)^{N-1-i} [\text{Pr}(d_{i+1} > r)G - C_i] \quad (12)$$

其中, x 是某个对等节点选择作为超级组对等节点 SGP 的概率, $\binom{N-1}{i} x^i (1-x)^{N-1-i}$ 是其他 $N-1$ 个对等节点中有 i 个对等节点选择作为超级组对等节点 SGP 的概率。

类似地, 当对等节点选择不作为超级组对等节点 SGP 时的平均收益可以通过式(13)进行计算

$$\bar{U}_{OP}(x) = \sum_{i=1}^{N-1} \binom{N-1}{i} x^i (1-x)^{N-1-i} \text{Pr}(d_i > r)G \quad (13)$$

根据式(12)和式(13), 对等节点的平均收益为

$$\bar{U}(x) = x\bar{U}_{SGP}(x) + (1-x)\bar{U}_{OP}(x) \quad (14)$$

将式(14)代入到复制者动态微分方程(11), 可得

$$\frac{dx}{dt} = x(1-x)[\bar{U}_{SGP}(x) - \bar{U}_{OP}(x)] \quad (15)$$

在均衡点 x^* , 没有参与者愿意偏离最优策略, 即有 $\frac{dx}{dt}|_{x^*} = 0$, 解该微分方程得 $x^* = 0, x^* = 1$, 或者 x^* 为方程 $\bar{U}_{SGP}(x) - \bar{U}_{OP}(x) = 0$ 的解。然而, 由于 $\frac{dx}{dt}|_{x^*} = 0$ 只是 x^* 成为演进稳定策略 ESS 的必要条

件, 进一步分析了每个演进稳定策略 ESS 候选策略的充分条件, 并得出如下结论, 并分别在定理 2、定理 3 和定理 4 中给出了相关证明。

1) 只有当 $\text{Pr}(d_1 > r)G - C > 0$ 时, $x^* = 0$ 是一个演进稳定策略 ESS。

2) 只有当 $\text{Pr}(d_N > r)G - \text{Pr}(d_{N-1} > r)G > C$ 时, $x^* = 1$ 是一个演进稳定策略 ESS。

3) 令 x^* 是方程 $\bar{U}_{SGP}(x) = \bar{U}_{OP}(x)$ 的解, 并且 $x^* \in (0, 1)$, 则 x^* 是一个演进稳定策略 ESS。

引理 1 令 $g(x) = \bar{U}_{SGP}(x) - \bar{U}_{OP}(x)$, 则有 $g'(x) < 0, \forall x \in [0, 1]$ 。

证明 根据式(12)和式(13), 有

$$g(x) = \sum_{i=0}^{N-1} \binom{N-1}{i} x^i (1-x)^{N-1-i} p_i - C \quad (16)$$

其中, $p_i = [\text{Pr}(d_{i+1} > r) - \text{Pr}(d_i > r)]G$, 表示在一个已经有 i 个超级组对等节点的对等组内再多引入一个超级组对等节点 SGP 所能够获得的额外收益。

对于 $\forall x \in [0, 1]$, 将 $g(x)$ 对 x 求导可得

$$\begin{aligned} g'(x) &= \sum_{i=0}^{N-1} \binom{N-1}{i} x^i (1-x)^{N-1-i} p_i - C \\ &= \sum_{i=0}^{N-1} \binom{N-1}{i} [ix^{i-1}(1-x)^{N-1-i} - x^i(N-1-i)(1-x)^{N-2-i}] p_i \\ &= \sum_{i=0}^{N-1} \binom{N-1}{i} x^{i-1} (1-x)^{N-2-i} (i(1-x) - x(N-1-i)) p_i \\ &= \sum_{i=0}^{N-1} \binom{N-1}{i} x^{i-1} (1-x)^{N-2-i} [i-x(N-1)] p_i \end{aligned} \quad (17)$$

引入一个整数 i_1 , 使得 $i_1 < (N-1)x$, 而 $i_1 + 1 > (N-1)x$, 这时有

$$\begin{aligned} g'(x) &= \sum_{i=0}^{i_1} \binom{N-1}{i} x^{i-1} (1-x)^{N-2-i} [i-x(N-1)] p_i + \\ &\quad \sum_{i=i_1+1}^{N-1} \binom{N-1}{i} x^{i-1} (1-x)^{N-2-i} [i-x(N-1)] p_i \end{aligned} \quad (18)$$

由于 p_i 是关于 i 的递减函数, 因此有: 当 $\forall i < i_1$ 时, $p_i > p_{i_1}$, 而对于 $\forall i > i_1$ 时, $p_i < p_{i_1}$ 。因此, 根据式(18), 有

$$g'(x) < \sum_{i=0}^{i_1} \binom{N-1}{i} x^{i-1} (1-x)^{N-2-i} [i-x(N-1)] p_i +$$

$$\begin{aligned}
& \sum_{i=i_i+1}^{N-1} \binom{N-1}{i} x^{i-1} (1-x)^{N-2-i} [i-x(N-1)] p_i \\
&= p_{i_i} \sum_{i=0}^{N-1} \binom{N-1}{i} x^{i-1} (1-x)^{N-2-i} [i-x(N-1)] \\
&= p_{i_i} \frac{d \left[\sum_{i=0}^{N-1} \binom{N-1}{i} x^i (1-x)^{N-1-i} \right]}{dx} = 0 \quad (19)
\end{aligned}$$

定理 2 $x^*=0$ 是演进稳定策略 ESS 的条件是 $\Pr(d_N = r)G - C = 0$ 。

证明 根据式(12)~式(14)，一个对等节点使用混合策略 x ，而其他对等节点使用混合策略 $x^*=0$ 时的平均收益为

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

其中， $\bar{U}_{SGP}(0) = \Pr(d_1 = r)G - C$ ， $\bar{U}_{OP}(0) = 0$ 。

1) 若 $\Pr(d_1 = r)G - C > 0$ ，即 $\bar{U}_{SGP}(0) > \bar{U}_{OP}(0)$ ，则每个对等节点都希望偏离到 $x=1$ ，以获取收益 $\bar{U}_{SGP}(0)$ ，而不是收益 $\bar{U}_{OP}(0)$ 。

2) 若 $\Pr(d_1 = r)G - C < 0$ ，即 $\bar{U}_{SGP}(0) < \bar{U}_{OP}(0)$ ，则每个对等节点都希望保持 $x=0$ ，以获取收益 $\bar{U}_{OP}(0)$ ，而不是收益 $\bar{U}_{SGP}(0)$ 。

3) 若 $\Pr(d_1 = r)G - C = 0$ ，即 $\bar{U}_{SGP}(0) = \bar{U}_{OP}(0)$ ，则 $g(0) = \bar{U}_{SGP}(0) - \bar{U}_{OP}(0) = 0$ 。根据引理 1，可知 $g'(x) < 0, \forall x \in [0, 1]$ ，因此有 $g(x) = \bar{U}_{SGP}(x) - \bar{U}_{OP}(x) < g(0) = 0, \forall x \in [0, 1]$ 。在这种情况下，有 $\bar{U}(0, x) = \bar{U}_{OP}(x) > \bar{U}(x, x) = \bar{U}_{OP}(x) + x(\bar{U}_{SGP}(x) - \bar{U}_{OP}(x))$ ，根据定义 1 可知， $x^*=0$ 是演进稳定策略 ESS。

因此，只有当 $\Pr(d_1 = r)G - C = 0$ 时， $x^*=0$ 是一个演进稳定策略 ESS。

定理 3 $x^*=1$ 是演进稳定策略 ESS 的条件； $\Pr(d_N = r)G - \Pr(d_{N-1} = r)G > C$ 。

证明 根据式(12)~式(14)，一个对等节点使用混合策略 x ，而其他对等节点使用混合策略 $x^*=1$ 时的平均收益为

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

其中， $\bar{U}_{SGP}(1) = \Pr(d_N = r)G - C$ ， $\bar{U}_{OP}(1) = \Pr(d_{N-1} = r)G$ 。

1) 若 $\Pr(d_N = r)G - \Pr(d_{N-1} = r)G < C$ ，即 $\bar{U}_{OP}(1) > \bar{U}_{SGP}(1)$ ，则每个对等节点都会偏离到 $x=0$ 以获得 $\bar{U}_{OP}(1)$ ，而不是 $\bar{U}_{SGP}(1)$ 。

2) 若 $\Pr(d_N = r)G - \Pr(d_{N-1} = r)G > C$ ，即 $\bar{U}_{OP}(1) > \bar{U}_{SGP}(1)$ ，则每个对等节点都会保持 $x=1$ 以获得 $\bar{U}_{SGP}(1)$ ，而不是 $\bar{U}_{OP}(1)$ 。

3) 若 $\Pr(d_N = r)G - \Pr(d_{N-1} = r)G = C$ ，即 $\bar{U}_{OP}(1) = \bar{U}_{SGP}(1)$ ，则 $\bar{U}(x, 1) = \bar{U}_N(1)$ ，并且有 $g(1) = \bar{U}_{SGP}(1) - \bar{U}_{OP}(1) = 0$ 。根据引理 1，可知 $g'(x) < 0, \forall x \in [0, 1]$ ，因此有 $g(x) = \bar{U}_{SGP}(x) - \bar{U}_{OP}(x) > g(1) = 0, \forall x \in [0, 1]$ 。此时，有 $\bar{U}(1, x) = \bar{U}_{OP}(x) + 1 \times (\bar{U}_{SGP}(x) - \bar{U}_{OP}(x)) > \bar{U}(x, x) = \bar{U}_{OP}(x) + x(\bar{U}_{SGP}(x) - \bar{U}_{OP}(x))$ ，由定义 1 可知， $x^*=1$ 是演进稳定策略 ESS。因此，只有当 $\Pr(d_N = r)G - \Pr(d_{N-1} = r)G = C$ 时， $x^*=0$ 是一个演进稳定策略 ESS。

定理 4 若 $x^* \in (0, 1)$ 是方程 $\bar{U}_{SGP}(x) = \bar{U}_{OP}(x)$ 的解，则 x^* 是一个演进稳定策略 ESS。

证明 令 $\bar{U}_i(x, x^*)$ 是对等节点 i 使用混合策略 x ，而其他对等节点使用混合策略 x^* 时对等节点 i 的平均收益。这时有

$$\bar{U}_i(x, x^*) = x\bar{U}_{SGP}(x^*) + (1-x)\bar{U}_{OP}(x^*) \quad (20)$$

由于 x^* 是方程 $\bar{U}_{SGP}(x) = \bar{U}_{OP}(x)$ 的一个解，有 $\bar{U}_{SGP}(x^*) = \bar{U}_{OP}(x^*)$ ，这时式(20)变为

$$\bar{U}_i(x, x^*) = \bar{U}_{SGP}(x^*) = \bar{U}_i(x^*, x^*) \quad (21)$$

这说明 x^* 满足定义 1 中的均衡条件。并且根据式(14)，有

$$\bar{U}_i(x, x) = \bar{U}_{OP}(x) + x(\bar{U}_{SGP}(x) - \bar{U}_{OP}(x)) \quad (22)$$

$$\bar{U}_i(x^*, x) = \bar{U}_{OP}(x) + x^*(\bar{U}_{SGP}(x) - \bar{U}_{OP}(x)) \quad (23)$$

因此，有

$$\bar{U}_i(x^*, x) - \bar{U}_i(x, x) = (x^* - x)(\bar{U}_{SGP}(x) - \bar{U}_{OP}(x)) \quad (24)$$

根据引理 1，知道 $g(x) = \bar{U}_{SGP}(x) - \bar{U}_{OP}(x)$ 是一个单调递减函数，由于 $\bar{U}_{SGP}(x^*) = \bar{U}_{OP}(x^*)$ ，当 $x < x^*$ 时， $\bar{U}_{SGP}(x) - \bar{U}_{OP}(x) > 0$ ；而当 $x > x^*$ 时， $\bar{U}_{SGP}(x) - \bar{U}_{OP}(x) < 0$ 。因此，对于 $\forall x \neq x^*$ 有 $(x^* - x)(\bar{U}_{SGP}(x) - \bar{U}_{OP}(x)) > 0$ ，即：

$$\bar{U}_i(x^*, x) > \bar{U}_i(x, x), \forall x \neq x^* \quad (25)$$

这说明 x^* 满足定义 1 中的稳定条件，根据式

(20)~式(25)可知 x^* 是一个演进稳定解 ESS。

5 分布式学习算法 ESS-SGP

通过求解复制者动态方程(11)来找到演进稳定策略 ESS，但这样做需要对等节点之间互相交换私有信息以及其他对等节点采取的策略。为此本文提出一种基于强化学习 Q-learning^[22]的分布式学习算法，使对等节点之间不需要交换私有信息就可以逐步地汇聚到演进稳定策略 ESS。

首先对复制者动态方程(11)进行离散化，可得如下形式

$$x_i(t+1) = x_i(t) + [\bar{U}_i(B, x_{-i}(t)) - \bar{U}_i(x_i(t))]x_i(t) \quad (26)$$

其中， t 为时间槽索引， $x_i(t)$ 为对等节点 i 在时间槽 t 内选择作为超级组对等节点 SGP 的概率。假定每一个时间槽可以被划分为 M 个子槽，每个对等节点可以在每个子槽的开始时刻选择是作为超级组对等节点 SGP，还是作为普通对等节点 OP。

为了计算 $\bar{U}_i(B, x_{-i}(t))$ 和 $\bar{U}_i(x_i(t))$ ，定义一个指示函数 $o_i(t, k)$ 。

定义 2 若对等节点 i 在时间槽 t 的第 k 个子槽开始时刻选择作为超级组对等节点 SGP，则指示函数 $o_i(t, k)$ 取值为 1；否则该函数取值为 0。

在时间槽 t 的第 k 个子槽，对等节点 i 的收益可以通过下式进行计算：

$$U_i(t, k) = \begin{cases} G - C_i, & \text{如果对等节点 } i \text{ 选择作为 SGP, 并且 } r^i > r \\ -C_i, & \text{如果对等节点 } i \text{ 选择作为 SGP, 并且 } r^i < r \\ G, & \text{如果对等节点 } i \text{ 选择作为 OP, 并且 } r^i > r \\ 0, & \text{如果对等节点 } i \text{ 选择作为 OP, 并且 } r^i < r \end{cases} \quad (27)$$

其中， r^i 是超级组对等节点 SGP 的总下载速率， r 为源速率。则 $\bar{U}_i(B, x_{-i}(t))$ 可以采用式(28)进行近似计算：

$$\bar{U}_i(B, x_{-i}(t)) = \frac{\sum_{k=1}^M U_i(t, k) o_i(t, k)}{\sum_{k=1}^M o_i(t, k)} \quad (28)$$

类似地， $\bar{U}_i(x_i(t))$ 可以采用下面的公式进行计算：

$$\bar{U}_i(x_i(t)) = \frac{1}{M} \sum_{k=1}^M U_i(t, k) \quad (29)$$

在分布式 Q-learning 算法中，使用 Q 值(即 $Q(t)$)来表示每个对等组关于未来收益的知识，并根据该知识来进行决策。分布式 SGP 选择算法的详细过程如算法 1 所示。

算法 1 分布式 SGP 选择算法 ESS-SGP

步骤 1 设时间槽索引 $t = 0$ ，每个对等节点使用随机产生的 $[0,1]$ 区间上的一个随机数 $x_i(0)$ 初始化自己的 x_i 值： $x_i(t) = x_i(0)$ 。

步骤 2 对等组中的所有对等节点初始化 Q 值为 $Q(t) = 0$ 。

步骤 3 loop //对对等组中的每个对等节点执行该循环

在时间槽 t 中，对于 $k = 1 : M$

if $x_i(t) > g$ **then**

对等节点 i 随机选择作为一个 SGP 从组外的 SGP 以下载速率 $r_i(t, k)$ 下载流媒体数据分片；或选择作为 OP 直接从组内 SGP 处下载流媒体数据分片

else

选择策略 $b^* = \arg \max_b Q(k)$

end if

对等节点 i 计算其指示函数 $o_i(t, k)$ 并使用式(27)计算其收益 $U_i(t, k)$ 。

步骤 4 对等节点 i 使用式(28)和式(29)计算其平均收益 $\bar{U}_i(B, x_{-i}(t))$ 和 $\bar{U}_i(x_i(t))$ 。

步骤 5 使用式(26)更新作为 SGP 的概率 $x_i(t)$ 。

步骤 6 更新 Q 值： $Q(k+1) = (1-l)Q(k) + l(\bar{U}_i(x_i(t)) + b \max_b Q(k))$ 。

步骤 7 endloop

在 ESS-SGP 选择算法中，对等节点首先使用概率 g 进行策略的随机选择，而 l 用来控制 Q 值的调整速度。根据上一轮迭代得到的 $Q(k)$ 和本轮计算得到的收益值 $\bar{U}_i(x_i(t))$ 可以计算得到新的 Q 值 $Q(k+1)$ ，它表示未来所期望的收益。

6 实验测试与分析

为了验证本文提出的基于演进博弈的 ESS-SGP 选择算法，在郑州大学校园中部署的 P2P 实时流媒体平台 LStream^[18] 上进行了算法测试。在实验过程中，设节点可获得的收益值 $G=1$ ， r^L 和 r^U 的值分别设定 50 和 1 000，假定在同类对等组内有 20 个对等节点，

它们的开销 C 为区间 $[0,0.2]$ 上的随机数。将本文方法 ESS-SGP 与传统的 P2P 流媒体系统中随机选择对等节点作为 SGP 的机制进行了实验比较,传统随机选择 SGP 机制方法记做 Random-SGP,在这类 P2P 实时流媒体网络中,如 PPLive^[23]和 Cool-Streaming^[24],每个对等节点从自己的伙伴列表中随机选择一些对等节点来进行流媒体数据片片的下载。

实验 1 比较了 ESS-SGP 算法与传统的 Random-SGP 方法在取得的社会福利(即所有对等节点的收益之和)方面的差别。如图 2 所示,ESS-SGP 算法可以获得一个稳定的纳什均衡解,通过对等节点之间的合作并且自适应地选择合适数目的超级组对等节点 SGP,ESS-SGP 算法在较高的视频源速率的情况下,仍然可以保持一个较高的社会福利水平,而传统 Random-SGP 方法的社会福利水平则随着源速率的增加呈线性快速下降趋势。

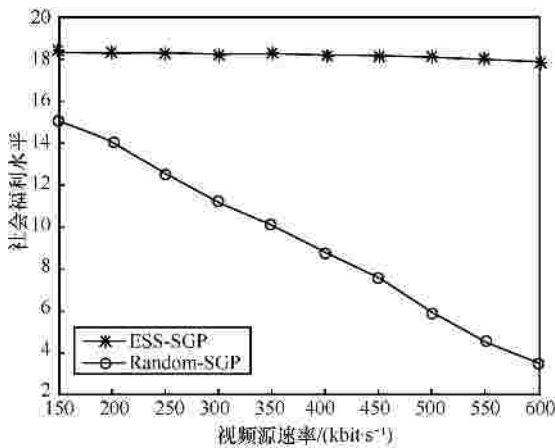


图 2 Random-SGP 方法与 ESS-SGP 方法取得的社会福利的比较

实验 2 对 ESS-SGP 算法的收敛性进行了评估。在视频源速率为 $r = 500 \text{ bit/s}$ 的情况下,当对等节点组中的同类对等节点的个数不同时,SGP 选择博弈的复制者动态性如图 3 所示。从图 3 中可以看出,所有对等节点在初始时选择作为 SGP 的概率都比较高,随着 P2P 流媒体系统运行时间的增长,所有对等节点逐渐地减低了作为 SGP 的概率,因为选择作为普通节点 OP 通常会为他们带来更高的收益。然而,由于选择作为 SGP 的概率过低可能会增加没有对等节点做 SGP 的可能性,因此选择作为 SGP 的概率将最终收敛到某个特定值,这个特定值的大小取决于

组内对等节点的个数。

实验 3 比较了 ESS-SGP 方法与传统 Random-SGP 方法的所有 SGP 节点平均上传速率。SGP 节点平均上传速率是所有超级组对等节点上传速率之和与超级组对等节点总数的比值,反映了 SGP 节点为系统提供的上传能力。从图 4 中可以看出,本文提出的 ESS-SGP 算法所产生的 SGP 节点为系统提供了更高的上传能力,且比较稳定,随着系统运行时间的增加波动性不大。

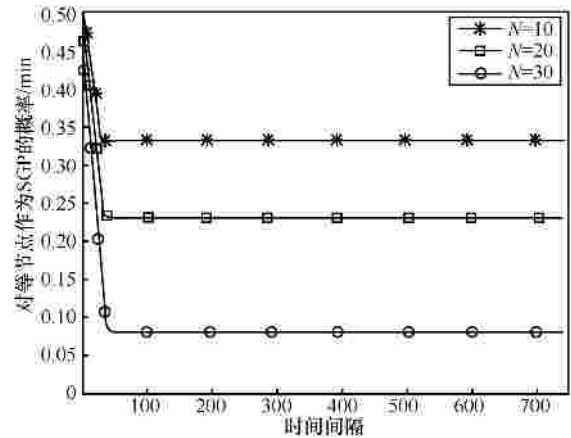


图 3 同构对等节点组中节点的行为动态性

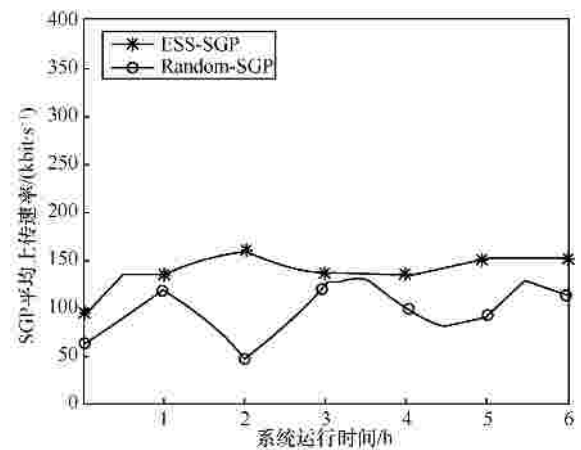


图 4 ESS-SGP 与 Random-SGP 方法在 SGP 平均上传速率方面的比较

在实验 4 中比较了 ESS-SGP 方法与传统 Random-SGP 方法在 P2P 流媒体系统的超级节点上传能力方面的差异。从图 5 中可以看出,没有引入超级组对等节点激励选择机制时,当 P2P 流媒体系统节点数达到 2000 个左右时,超级节点的负载接近饱和;当节点规模继续增加时,系统的上传能力增加不明显。引入 ESS-SGP 选择机制后,消除了系统瓶颈,P2P 网络的上传能力随节点增加而稳定增加。

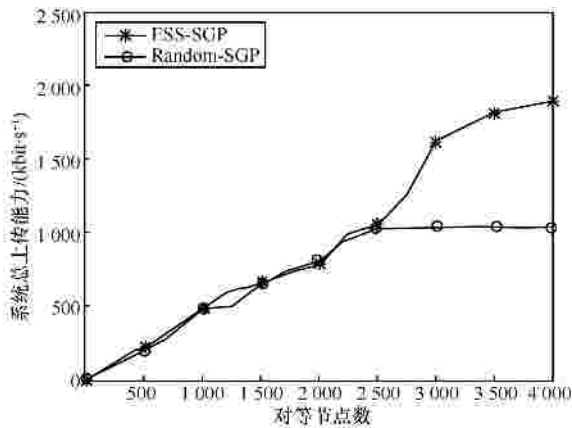


图 5 ESS-SGP 与 Random-SGP 方法在 P2P 流媒体系统网络上传能力方面的对比

7 结束语

针对分层混合 P2P 实时流媒体系统中的超级节点选择问题,本文提出了一种用于激励对等组内合作的 SGP 节点选择博弈模型。将位于分层混合 P2P 流媒体系统边缘的对等节点按照地域位置划分为不同的对等组,并针对组内对等节点行为的自私性和理性特点,使用演进博弈论框架对 SGP 选择行为进行了建模,促进并鼓励组内对等节点之间的合作。其次,为每个对等节点推导出了演进稳定策略 ESS,提出了一种不需要对等节点间交换私有信息就可以逐步地汇聚到演进稳定策略的分布式学习算法 ESS-SGP。从实验结果可以看出,与传统的随机选择超级节点机制的 P2P 实时流媒体方案相比,本文提出的 ESS-SGP 算法所选择产生的超级组对等节点能够稳定地为系统提供更高的流媒体上传速率,能够使整个 P2P 系统的上传能力随着对等节点数目的增加而稳定增加,并能够为每个对等节点带来更好的收益。

本文没有讨论对等节点上有缓冲区的情况,同时假定了对等组内的对等节点都是同构的。新的协议增加了部分额外计算开销,因此还存在一定的局限性。进一步的研究工作将考虑对等组内的对等节点异构的情况,使之更贴近于 P2P 实时流媒体系统的实际情况。

参考文献:

[1] RAMZAR N, PARK H, IZQUIERDO E. Video streaming over P2P networks: challenges and opportunities[J]. *Signal Processing: Image Communication*, 2012, 27(5): 401-411.

[2] LO V, ZHOU D, LIU Y, *et al.* Scalable supernode selection in

peer-to-peer overlay networks[A]. *Proceedings of Second International Workshop on Hot Topics in Peer-to-Peer Systems 2005*[C]. San Diego, CA, 2005.18-25.

[3] MIN S, CHO D. Super-peer selection based framework using dynamic capacity and similarity[A]. *Proceedings of 21th International Symposium on Computer and Information Sciences-ISCIS 2006*[C]. Springer Berlin Heidelberg, 2006.803-812.

[4] LU G, MIN P. Optimal superpeer selection based on load balance for P2P file-sharing system[A]. *Proceedings of International Joint Conference on Artificial Intelligence*, 2009[C]. Los Angeles, USA, 2009. 92-95.

[5] MIN S H, HOLLIDAY J, CHO D S. Optimal super-peer selection for large-scale P2P system[A]. *Proceedings of IEEE International Conference on Hybrid Information Technology*, 2006[C]. Cheju Island, Korea 2006.588-593.

[6] JIN Y, ZHAO H. Trust-based supernode selection in peer-to-peer systems[A]. *Proceedings of 2nd International Conference on Future Computer and Communication (ICFCC 2010)*[C]. Wuhan, Chi 2010. 285-289.

[7] YU C, GOU X, ZHANG C, *et al.* Supernode election algorithm in P2P network based upon district partition[J]. *International Journal of Digital Content Technology and Its Applications*, 2011, 5(1): 186.

[8] GAO Z, GU Z, WANG Z, *et al.* SPSI: A hybrid super-node election method based on information theory[A]. *Proceedings of IEEE 14th International Conference on Advanced Communication Technology (ICACT 2012)*[C]. Seoul, Korea, 2012.1076-1081.

[9] WOLF S. On the Complexity of the Incapacitated Single Allocation p -hub Median Problem with Equal Weights[R]. University of Kaiserslautern, 2007.

[10] WOLF S, MERZ P. Evolutionary local search for the super-peer selection problem and the p -hub median problem[J]. *Hybrid Metaheuristics Lecture Notes in Computer Science*, 2007, 4771: 15.

[11] KRATICA J, KOJIC J, TOŠIC D, *et al.* Two hybrid genetic algorithms for solving the super-peer selection problem[J]. *Advances in Intelligent and Soft Computing*, 2009, 58: 337-346.

[12] WANG S L, YANG S, SHEN K, *et al.* A super node selecting mechanism based on AHP[A]. *Proceedings of IEEE Seventh International Conference on Grid and Cooperative Computing*, 2008. (GCC'08)[C]. Shenzhen, China, 2008. 403-406.

[13] SHOHAM Y. Computer science and game theory[J]. *Communication of the ACM*, 2008, 51(8): 74-79.

[14] ROY S, ELLIS C, SHIVA S, *et al.* A survey of game theory as applied to network security[A]. *Proceedings of 43rd Hawaii International Conference on System Sciences*[C]. Poipu, Hawaii, 2010.1-10.

[15] CHISCI L, PAPI F, PECORELLA T, *et al.* An evolutionary game

approach to P2P video streaming[A]. Proceedings of IEEE Global Telecommunications Conference (GLOBECOM 2009)[C]. Hawaii, USA, 2009.1-5.

- [16] BIOGLIO V, GAETA R, GRANGETTO M, *et al.* A game theory framework for ISP streaming traffic management[J]. Performance Evaluation, 2011, 68(11): 1162-1174.
- [17] LStream P2P live streaming system[EB/OL]. <http://lstream.ha.edu.cn/>.
- [18] LV X, WU Q, LI R, *et al.* On tree construction of super peers for hybrid P2P live media streaming[A]. Proceedings of 19th International Conference on Computer Communications and Networks (ICCCN 2010)[C]. ETH Zurich, Switzerland, 2010. 1-6.
- [19] HUANG Y, FU T, CHIU D M, *et al.* Challenges, design and analysis of a large-scale P2P-vod system[J]. ACM SIGCOMM Computer Communication Review, 2008, 38(4): 375-388.
- [20] SMITH J. Evolutionary and the Theory of Games[M]. Cambridge University Press, 1982.
- [21] WANG B, LIU K, CLANCY T. Evolutionary cooperative spectrum sensing game: how to collaborate[J]. IEEE Trans Commun, 2010, 58(3): 890-900.
- [22] FUDENBERG D, LEVINE D. The Theory of Learning in Games[M]. Cambridge: MIT Press, 1998.
- [23] SPOTO S, GAETA R, GRANGETTO M, *et al.* Analysis of PPLive through active and passive measurements[A]. Proceedings of IEEE International Symposium on Parallel and Distributed Processing 2009(IPDPS 2009)[C]. 2009. 1-7.
- [24] LI B, XIE S, QU Y, *et al.* Inside the new coolstreaming: principles, measurements and performance implications[A]. Proceedings of 27th IEEE Conference on Computer Communications (INFOCOM 2008)[C]. Phoenix, AZ, USA, 2008. 1031-1039.

作者简介：



陈静 (1977-), 女, 河南禹州人, 西安交通大学博士生, 郑州大学讲师, 主要研究方向为 P2P 流媒体、博弈论和网络安全。



周兵 (1964-), 男, 河南永城人, 郑州大学教授, 主要研究方向为多媒体信息处理及传输、智能视频监控、图像/视频处理等。



任晨珊 (1991-), 女, 河南郑州人, 郑州大学硕士生, 主要研究方向为多媒体通信。



董小社 (1963-), 男, 陕西西安人, 西安交通大学教授、博士生导师, 主要研究方向为分布式计算、网络安全等。