

# P2P 流媒体系统中基于直觉模糊集 的节点选择策略

李彦 王丽娜

(重庆理工大学计算机科学与工程学院 重庆 400054)

**摘要** 通过对非结构化 P2P 网络中的随机节点选择算法进行研究,运用模糊理论中的直觉模糊集概念提出了基于得分函数的节点选择策略,以克服传统策略的随机性和盲目性。采取对节点的可靠度进行得分排序的方法,充分利用了能力强的节点,其具有良好的适应性,可反映 P2P 网络中节点和资源的动态性,提高了系统的整体效率。

**关键词** P2P 流媒体,节点选择,直觉模糊集,得分函数

**中图分类号** TP3-0 **文献标识码** A

## Node Selection Strategy Based on Intuitionistic Fuzzy Sets in P2P Streaming Media System

LI Yan WANG Li-na

(School of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China)

**Abstract** Through studying random node selection algorithm of unstructured P2P networks, By applying the theory of intuitionistic fuzzy set and put forward a node selection strategy which based on the score function, in order to overcome the randomness and blindness in the traditional strategy. This paper adopt scoring method of nodes reliability, make full use of good nodes, It has good adaptability, and can reflect the node and resources dynamic in P2P network, Meanwhile it improve the overall efficiency of the system.

**Keywords** P2P streaming media, Node selection, Intuitionistic fuzzy sets, Scoring function

### 1 引言

P2P 对等网络模型改变了传统互联网 C/S 模型的通信和业务模式<sup>[1]</sup>,让更多终端能够参与到互联网应用中。而 P2P 应用导致用户间流量激增,用户上下行接入带宽均被高负荷使用,互联网传统的 C/S 流量模型基本被改变。由于 P2P 网络本身的可扩展性,基于 P2P 方式的流媒体技术很好地解决了传统流媒体带宽不足的问题。典型的 P2P 流媒体应用有: PPLive, PPStream, Gridcast 等。

相比传统的 C/S 模式, P2P 流媒体系统具有无法比拟的优势。然而想充分利用 P2P 网络上的信息资源,就要对网络上的资源进行搜索,资源搜索的关键是对网络节点的选择。文献[2]提出了固定随机邻居节点选择算法 FRNS。在动态环境中,该算法通过消息传递机制查找可能成为邻居节点的在线节点,一旦找到,就把此在线节点放到候选邻居列表中。列表中的候选节点在成为邻居节点之前,如果该节点到了老化时间,将会从列表中删除。如果候选列表中的在线节点数大于零,且请求节点的邻居数小于最大邻居数,就在列表中随机选择邻居节点,直到满足固定邻居数。由于随机选择算法没有考虑到节点存储资源能力的差异性,且无法保证节点的服务质量,选出的节点所构造的逻辑网络可能与实际物理网络不匹配,造成覆盖网压力过大<sup>[3]</sup>。虽然随机选择能加强系统弹性,但这样也带来了较严重的数据冗余,不适合高带宽需

求的流媒体应用,并且它的随机性使得搜索的效率变低。

本文综合考虑了 P2P 网络中影响节点选择的多种因素,提出了一种改进的节点选择算法 HSA。选择基于 FRNS 节点选择算法做改进,因为它的协议简单,能够保证系统的鲁棒性和负载均衡。下面从节点选择、模糊理论及基于直觉模糊集的节点选择策略 HAS 方面进行详细阐述,并通过相关分析,得出该算法在理论上能有效地提高 P2P 网络中节点选择的效率。

### 2 节点选择策略

在 P2P 流媒体系统中,把节点和路由器在物理网络层面上的相互连接称为底层网络(underlay),把节点之间的逻辑连接称为覆盖网(overlay)<sup>[4]</sup>。所谓的节点,就是参与系统的终端,一般包括媒体源服务器、普通客户端和转发服务器。节点选择机制是指:在 P2P 系统中,当一个节点申请加入该系统时,系统中已经存在了一组具有该节点所需内容的其它节点,系统将根据某种决策来选取其中的部分节点作为该节点的邻居节点,随后此节点将与这些邻居节点建立网络连接,从而获取数据。如图 1 所示:有一覆盖网,当节点 D 请求加入时,节点选择服务器根据某种策略选择了节点 A、C、E 作为它的邻居节点,然后进行数据传输。

节点选择算法是混合式拓扑结构的 P2P 流媒体系统中的核心技术之一,当一个节点加入时,有多个节点可以为它提

本文受重庆市教委科技基金项目(KJ120825)资助。

李彦(1976-),男,博士生,副教授,主要研究方向为计算机网络与软件, E-mail: ly@cqu.edu.cn; 王丽娜(1987-),女,主要研究方向为计算机网络与软件。

供资源,选择哪些节点为它服务是 P2P 流媒体服务体系中需要解决的最首要问题。经过对大量系统的调查研究,将典型的节点选择算法划分为随机选择算法、带宽优先选择算法和拓扑聚焦算法<sup>[5]</sup>。

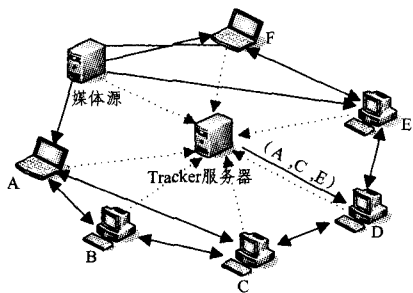


图1 节点选择机制结构

邻居节点的选择可以根据不同的 QoS 需求采取不同的策略。如果需要流媒体的服务延迟较小,可以选择物理临近的节点建立会话,即尽量选择同一 ISP 内的节点。如果希望高质量的流媒体服务,可以选择高带宽、CPU 能力强的节点,如:在宽带接入的 PC 机和 ADSL 接入的终端中选择前者。如果要求比较稳定的流媒体服务,则应选择相对稳定的节点,如在系统中停留时间较长,则不会频繁加入和退出系统的节点。一般情况下,节点的选择是多种策略的折中,本文提出的 HAS 节点选择算法就是根据不同的 QoS 需求,综合考虑最常见的用户需求,基于节点的得分来确定一个节点是相对可靠的,这需要该节点的一些历史数据,如在线时间、节点贡献度等信息。

### 3 基于直觉模糊集的节点选择算法实现

Zadeh<sup>[6]</sup>于1965年提出的模糊集理论已经较广泛地应用于模糊决策之中。模糊决策就是指在不确定性环境下进行决策的数学理论及方法。事实上,现实世界中的决策大多是模糊决策。为了更好地处理不确定信息,Atanassov<sup>[7]</sup>于1983年提出了直觉模糊集的概念,P2P流媒体系统中的节点选择是以选择最可靠的节点作为邻居节点为目标,这里的“最可靠”就是一个模糊概念。基于此我们就用直觉模糊集对可靠节点的选择做最优决策,给出基于得分函数的节点选择策略。

#### 3.1 直觉模糊集理论

在一个直觉模糊集中,用一个隶属函数  $U_A$  和一个非隶属函数  $V_A$  来描述其隶属度的边界,一个对象的支持度、反对度和犹豫度分别用  $U_A$ 、 $V_A$  和  $1-U_A-V_A$  来表示,这使得直觉模糊集在处理不确定的信息时比传统的模糊集理论有更强的表现能力以及灵活性。

直觉模糊集定义<sup>[8]</sup>:

设  $X=\{x_1, x_2, \dots, x_n\}$  是给定的论域,则  $X$  上的一个直觉模糊集  $A$  具有下列形式:

$$A = \{ \langle x_i, U_A(x_i), V_A(x_i) \rangle \mid x_i \in X \} \quad (1)$$

其中,  $U_A: X \rightarrow [0, 1]$ ,  $V_A: X \rightarrow [0, 1]$ , 且  $0 \leq U_A(x) + V_A(x) \leq 1, \forall x \in X$ ,  $U_A$  表示元素  $x$  对集合  $A$  的隶属度,  $V_A$  表示元素  $x$  对集合  $A$  的非隶属度。

$$\pi(x_i) = 1 - U_A(x_i) - V_A(x_i) \quad (2)$$

$\pi$  称为  $A$  的直觉指数,它是  $x$  对  $A$  的犹豫程度的一种测度。

#### 3.2 基于直觉模糊集的节点选择策略

作为模糊集的一种推广,直觉模糊集的特点就是同时考

虑隶属度、非隶属度的影响。这就使得直觉模糊集在处理带有模糊性的问题时更为灵活实用。文献<sup>[8]</sup>中提出的直觉模糊集可以用来在多属性的方案中做最优决策,考虑到 P2P 系统中的 Peer 节点不同于一般的服务器,它具有高度的动态性和异构性,每个节点可以用几个不同的属性来描述。因此,本文就致力于将直觉模糊集应用到对等网络中的节点选择,以提高 P2P 流媒体的 QoS 保障。

##### 3.2.1 算法原理

给出一个记分函数法:设有一个方案集  $A = \{A_1, A_2, \dots, A_m\}$ , 每个方案需要考虑  $n$  个属性,记属性集为  $C = \{C_1, C_2, \dots, C_n\}$ , 令  $U_{ij}, V_{ij}$  分别表示方案  $A_i$  对于属性  $C_j$  的满意和满意程度,且  $0 \leq U_{ij}, V_{ij} \leq 1$  及  $U_{ij} + V_{ij} \leq 1$ , 对属性  $C_1, C_2, \dots, C_n$  按其重要性进行赋权,设其权重分别是  $W_1, W_2, \dots, W_n$ , 其中  $W_1, W_2, \dots, W_n \in [0, 1]$  且  $W_1 + W_2 + \dots + W_n = 1$ 。可以用以下的决策矩阵  $R$  表示:

$$R = \begin{bmatrix} (U_{11}, V_{11}) & (U_{12}, V_{12}) & \dots & (U_{1n}, V_{1n}) \\ (U_{21}, V_{21}) & (U_{22}, V_{22}) & \dots & (U_{2n}, V_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ (U_{m1}, V_{m1}) & (U_{m2}, V_{m2}) & \dots & (U_{mn}, V_{mn}) \end{bmatrix} \quad (3)$$

决策者要在  $m$  个方案集中选出一个最优方案。在进行决策的过程中,为了表示决策者对方案  $A_i$  的满足程度,我们使用式(4)表示的记分函数<sup>[8]</sup>  $S$ :

$$S(A) = U_A - V_A \quad (4)$$

根据记分函数,方案  $A_i$  对属性  $C_j$  的满足程度可由下面的公式表示出来:

$$S(A_{ij}) = U_{ij} - V_{ij} \quad (5)$$

其中,  $S(A_{ij}) \in [-1, 1]$ 。要考虑属性的权重,决策者对方案  $A_i$  的满足程度可由下列公式表示:

$$g(A_i) = S(A_{i1})W_1 + S(A_{i2})W_2 + \dots + S(A_{in})W_n \quad (6)$$

其中,  $i=1, 2, \dots, m$ , 可以看出  $S(A_i)$  越大,方案  $A_i$  越满足决策者的要求。

##### 3.2.2 算法描述

考虑覆盖网中有  $n$  个候选节点且每个节点具有  $m$  个属性,通过模拟实验的测试和分析,最终选取网络带宽、节点贡献度、节点在线时间 3 个属性来考察某一节点是否为可靠性高的节点,其中  $bandwidth$  代表节点的网络带宽,  $contribution$  代表节点贡献度,  $online$  代表节点在线的时间。公式描述如下<sup>[9]</sup>。

节点的网络带宽:

$$bandwidth = bandwidth / \max\{bandwidth\} \quad (7)$$

即  $bandwidth$  为节点带宽与网络最大带宽的比值,其值在  $[0, 1]$  之间。

节点贡献度:

$$contribution = Bout / Bin \quad (8)$$

其中,  $Bout$  表示节点的输出带宽,  $Bin$  表示节点的输入带宽。

节点在线时间:

$$online = \text{节点平均一天在线时间} / 24 \quad (9)$$

将此转化为直觉模糊集环境下的多属性决策问题,即:论域  $X = \{x_1, x_2, \dots, x_n\}$ , 其中  $x_i$  代表覆盖网络中的节点  $i$ , 每个节点的属性集  $R = \{r_1, r_2, r_3\}$ , 分别代表网络带宽、节点贡献度、节点在线时间。设  $U_{ij}$  和  $V_{ij}$  分别表示节点  $x_i$  相对于属性

$r_i$  的隶属度和非隶属度。节点  $x_i$  可以用一个直觉模糊集表示为:

$$x_i = \{(r_1, U_{i1}, V_{i1}), (r_2, U_{i2}, V_{i2}), \dots, (r_3, U_{i3}, V_{i3})\}$$

其中,  $1 \leq i \leq n$ 。

在此可以用上面定义的计分函数  $S$  来计算节点  $x_i$  的得分  $g(x_i)$ , 然后得到“可靠性”较好的候选节点, 再结合节点的 IP 地址选择同一 ISP 范围内的最优节点作为邻居节点。

### 3.2.3 伪代码实现

Joining Node n :

Sending join message to node selection sever S;

Node selection sever S :

1. Creating candidate neighbor nodes list  $N = \{n_1, n_2, \dots, n_m\}$ ;
2. if  $0 < m \leq k$
3. then node n join in the overlay network;
4. else
5. goto step 6;
6. According to the history information of node, created a decision matrix  $R = [(U_{ij}, V_{ij})]$ ;
7. for ( $i=0; i < n; i++$ )
8. for ( $j=0; j < 3; j++$ )
9.  $S(A_{ij}) = U_{ij} - V_{ij}$ ;
10. endfor
11.  $g(n_i) = \sum_{j=0}^2 S(A_{ij}) * W_j$ ;
12. endfor
13. Reverse arrangement g, then taking first k node as neighbor nodes

### 3.3 算法的流程实现

定义节点的数据结构为 Node( Lnumber, ID, Group, Parents, Children, Ltime, Status, Grade, Prepare)。其中, Lnumber 是对节点排序后的编号; ID 为节点标识符; Group 表示根据节点的 IP 地址所得的节点分组; Parents 即为提供资源的节点; Children 表示接收资源的节点; Ltime 表示节点最后一次更新的时刻; Status 表示当前节点的状态; Grade 是该节点的经  $S$  函数计算后的得分; Prepare 表示备用节点信息<sup>[11,12]</sup>。

(1) Tracker 服务器<sup>[10]</sup> 即流媒体系统中用来管理节点、构建和记录网络拓扑结构的服务器。对覆盖网中所有的节点根据其 IP 地址进行分组, 定时对节点状态、“可靠度”得分进行实时记录和排序, 并保存在列表 Nodelist 中。

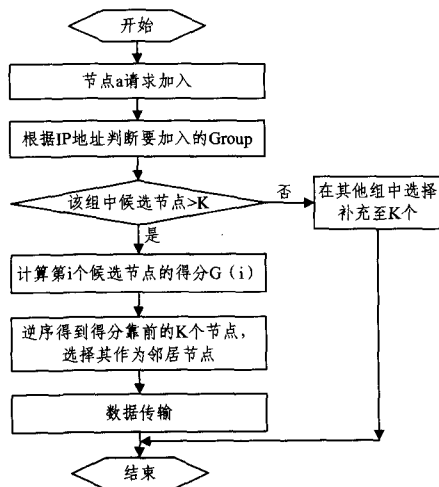


图2 算法流程图

(2) 如图 2 中所示, 当新节点申请加入覆盖网时, Tracker 服务器将该节点加入到 Nodelist 列表中, 并及时更新列表中的得分排序, 优先选择与该节点在同一 Group 中的节点, 其次选择其它组中的节点, 它们都是将 Grade 排序靠前的  $k$  个节点作为该节点的 Parents<sup>[12,13]</sup>。这里的 Tracker 服务器将及时更新该节点的上游节点信息。

(3) 列表中的节点收到来自其它节点的请求时, 先查找自己是否符合要求(即查看 Group 和 Grade 的信息), 若符合条件则对 Tracker 服务器反馈接收信息, Tracker 服务器收到来自该节点的接收信息时, 修改 Nodelist 列表中的 Parents 和 Children 信息, 并将该节点作为自己的 Children 节点, 开始向此节点传送媒体流; 否则, 向 Tracker 服务器返回拒绝信息。

(4) 如果 Tracker 服务器收到了来自该节点的拒绝信息或者在规定时间内没有收到来自该节点的信息, 将会继续执行步骤(3), 直到遍历系统中所有的节点为止<sup>[14]</sup>。

(5) 当节点申请退出系统时, 节点会分别向 Tracker 服务器和其上、下游节点发送请求, Tracker 服务器收到该请求后自动删除与该节点相关的信息, 而上、下游节点则会启用备用节点或者重新向服务器发出请求。

**结束语** 本文对 P2P 网络中广泛应用的随机节点选择算法进行研究, 提出基于直觉模糊集中得分函数的节点选择策略。HSA 节点选择算法在总体上优于随机选择算法。在运用模糊理论的 HSA 节点选择算法中, 选择得分比较高的可靠节点, 能保持较高的搜索效率。其主要原因在于: 在定义可靠性节点过程中, 综合考虑了节点提供的服务质量、节点贡献度等因素。因此, 选择高得分节点, 一方面提高了系统的安全性, 另一方面节点自身的带宽、在线时间等因素对节点选择的不良影响平均最小, 从一定程度上减小了系统的传输代价。

HSA 算法适用于大规模、节点性能低的环境, 并且具有简单、易于实现的优势。而对于节点性能比较好的情况, HSA 算法的优势不明显, 因为随机选择节点选择算法是基于冗余和随机这一主动机制来克服潜在的进程失效和链路失效的。本文只做了算法的理论研究, 该算法还有待在通用的 P2P 网络中进行检验。

### 参考文献

- [1] 张颖, 周娅, 黄桂敏, 等. 一种对等网络分层管理资源定位模型[J]. 计算机工程与应用, 2011, 47(19): 113-117
- [2] Mruthy S, Lunes G, Aceves J. An efficient routing protocol for wireless networks[J]. ACM Balzer Mobile Networks and Applications Journal, 1996(2): 183-197
- [3] 陈瑞昭, 刘永广. 基于能力因子的 P2P 邻居节点随机可变选择算法[J]. 计算机应用, 2010
- [4] 韦建楠, 庄雷. P2P 流媒体中未知覆盖网拓扑信息的节点选择策略[J]. 计算机应用研究, 2012
- [5] 何海亮, 黄道颖, 等. 两段式 P2P 流媒体节点选择算法[J]. 计算机工程, 2011
- [6] Zadeh L A. Fuzzy set[J]. Information and control, 1965, 8(3): 338-356
- [7] Atanassov K T. Intuitionistic fuzzy sets[J]. Fuzzy Sets and System, 1986, 20(1): 87-96

(下转第 319 页)

法识别图 2 中间接跳转目标地址,识别结果如图 5 所示。

其中图 5(a)所示的控制流图是按照文献[6]介绍的基于抽象解释的控制流图重构方法得到的控制流图,对于间接跳转,该方法把跳转目标表达式的所有值都作为间接跳转的目标地址,控制流图中的虚线表示所有可能的边;在文献[7]中,作者对上述方法进行了改进,对于间接跳转,直接对跳转目标赋值,按照此方法,在地址 7,令跳转目标  $x=14$ ,控制流图如图 5(b)所示。

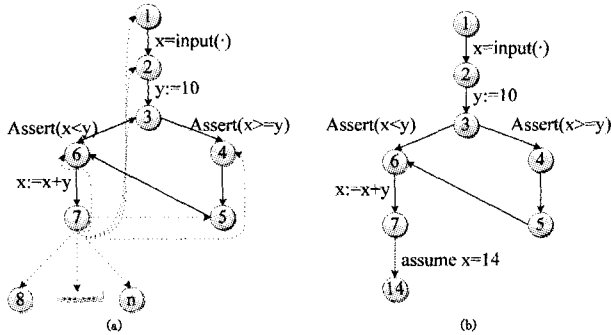


图 5 基于抽象解释的控制流图

通过对比发现,对于间接跳转目标地址的识别,基于抽象解释的控制流图重构引入了过多的目标地址,得到的控制流图不精确;改进后的方法虽然不会引入错误的跳转目标地址,但对于每个间接跳转只能得到一个目标地址,覆盖率低。而本文提出的基于路径约束的间接跳转目标识别方法可以依据从入口点到间接跳转的路径集合识别出多个目标地址,而且可以保证目标地址的正确性。

**结束语** 本文提出的基于路径约束的间接跳转目标识别方法充分利用了初始控制流图中隐含的约束条件和程序的动态执行信息对间接跳转的目标地址进行识别,相对于单纯的静态或动态分析方法,该方法较好地结合了两者的优点,可以有效识别出间接跳转的目标地址。

针对间接跳转的目标地址识别,当前缺乏一种通用的且能够精确识别其跳转目标地址的方法。本文提出的方法虽然能够准确识别出一个跳转目标地址的集合,但仍存在一些不足。首先,利用本文提出的方法得到的间接跳转目标地址集合是实际目标地址集合的一个子集,其虽然可以保证正确性,但在覆盖率上还有待提高,尤其是对于 switch-case 的跳转表形式。其次,该方法在工程实现上有所欠缺。下一步,计划对该方法进一步优化,充分利用程序的数据流信息和动态运行信息,在保证间接跳转目标地址识别正确性的同时,尽可能提高完整性。

## 参考文献

- [1] Balakrishnan G. WYSINWYX: What you see is not what you execute[D]. Wisconsin: University of Wisconsin, 2007
- [2] Cifuentes C. Reverse Compilation Techniques [D]. Queensland: Queensland University of Technology, 1994
- [3] Bardin S, Herrmann P, Veldrine F. Refinement-Based CFG Reconstruction from Unstructured Program [C] // Verification Model Checking and Abstract Interpretation (VMCAI), 2011. Berlin: LNCS 6538, 2011: 54-69
- [4] Larus J R, Schnarr E. EEL: Machine-Independent Executable Editing [C] // Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation, 1995: 291-300
- [5] De Sutter B, De Bus B, De Bosschere K. Link-time binary rewriting techniques for program compaction[J]. ACM Trans. Program. Lang. Syst., 2005, 27(5): 882-945
- [6] Kinder J, Zuleger F, Veith H. An Abstract Interpretation-Based Framework for Control Flow Reconstruction from Binaries [C] // Verification Model Checking and Abstract Interpretation (VMCAI), 2009. Berlin: LNCS 5043, 2009: 214-228
- [7] Kinder J, Kravchenko D. Alternating Control Flow Reconstruction [C] // Verification Model Checking and Abstract Interpretation (VMCAI), 2012. Berlin: LNCS 7148, 2012: 267-282
- [8] Xu L, Sun F, Su Z. Constructing precise control flow graphs from binaries [R]. Tech. Rep. University of California, Davis, 2009
- [9] Binary analysis platform (BAP) [OL]. <http://bap.ece.cmu.edu>, 2012-11-25
- [10] Song D, Brumley D, Yin Heng, et al. BitBlaze: A New Approach to Computer Security via Binary Analysis [C] // Intelligent Computing and Integrated Systems Security (ICISS), 2008. Berlin: LNCS 5352, 2008: 1-25
- [11] STP: A Decision Procedure for Bitvectors and Arrays [OL]. [http://people.csail.mit.edu/vganesh/STP\\_files/stp.html](http://people.csail.mit.edu/vganesh/STP_files/stp.html), 2012-11-01
- [12] 方震. 代码逆向分析中的语句恢复与算法识别技术研究 [D]. 郑州: 信息工程大学, 2009
- [13] 殷文建. 面向 ARM 体系结构的代码逆向分析关键技术研究 [D]. 郑州: 信息工程大学, 2009
- [14] Schwartz E J, Avgerinos T, Brumley D. All You Ever Wanted to Know About Dynamic Taint Analysis and Forward Symbolic Execution [C] // IEEE Symposium on Security and Privacy, 2010. Pittsburgh: DOI 10. 1109/SP, 2010, 26

(上接第 282 页)

- [8] 黎华, 王周敬. 基于直觉模糊集的多属性决策问题 [J]. 昆明理工大学学报: 理工版, 2008
- [9] 张妮妮, 范训礼, 等. 基于模糊理论的 P2P 流媒体节点选择算法 [J]. 计算机工程, 2009
- [10] 张治斌, 冯文峰, 黄永峰. 基于 Gossip 的自适应成员关系管理协议 [J]. 计算机应用, 2009(11): 2932-2935
- [11] 叶枫, 张思发. 基于 Gossip 协议的 P2P 流媒体直播系统的研究

[J]. 计算机与数字工程, 2009, 37(5): 89-93

- [12] 伍红华. 基于 Gossip 协议的 P2P 内容分发系统模型 [J]. 湖北教育学院学报, 2007, 24(8): 29-31
- [13] 袁雪萍, 周芳, 陈璐. 基于 Gossip 协议的 P2P 流媒体算法优化 [J]. 计算机与现代化, 2010(10): 130-141
- [14] 郭良敏, 杨寿保, 郭磊涛, 等. P2P 网络中基于区域划分的超级节点选取机制 [J]. 小型微型计算机系统, 2008, 29(2): 208-212