

云计算在轨道交通线网信息分发中的应用

韩泉叶^{1,2}, 王晓明², 党建武²

(1. 陕西广播电视大学 陕西工商职业学院 计算机信息管理系, 陕西 西安 710119;

2. 兰州交通大学 电子与信息工程学院, 甘肃 兰州 730070)

摘要:对城轨线网的数据信息及其特征进行了分析,讨论了推送和拉取的信息分发机制,提出了一种适合于城轨线网的信息分发机制,该机制将推送与拉取相结合,根据当前系统分发信息的特征和及时性要求,通过动态分配可用资源来确定推送和拉取,达到最小化数据块分发延时的目的。实验结果表明,推拉相结合的信息分发机制通过动态调节参数,既可适用于线网中严格一致性需求的信息流,也可适用于线网中非严格一致性需求的信息流,还可适用于线网中根据需要随时获取的信息流。

关键词:云计算;城轨线网;信息分发

中图分类号:TP311

文献标识码:A

文章编号:1673-629X(2012)12-0203-04

Application of Cloud Computing on Information Dissemination in Urban Rail Transit Line Network

HAN Quan-ye^{1,2}, WANG Xiao-ming², DANG Jian-wu²

(1. Department of Computer Information Management, Shaanxi Commerce and Vocational College,

Shaanxi Radio and TV University, Xi'an 710119, China;

2. Electronic and Information Engineering College, Lanzhou Jiaotong University,
Lanzhou 730070, China)

Abstract: The characteristics of data information in urban rail transit line network are analyzed. The push and pull information dissemination mechanism is discussed. In this paper, a kind of information disseminating mechanism is presented, which combines push approach with pull approach and fits for urban rail transit line network, by which the system can dynamically and intelligently allocate the available resource between push clients and pull clients so as to optimize the sending time delay of data chunk given their characters and desired currency. And the experimental results are presented to validate the approach. The major conclusion is that the approach of combination push with pull can fit for both stringent coherency requirements for data item and less stringent coherency requirements for data item on line network, also can fit for acquiring at any time according to requirement on line network, by optimizing the parameter dynamically.

Key words: cloud computing; urban rail transit line network; information dissemination

0 引言

云的构建技术是一种利用大量并行工作的处理器来处理高性能计算问题的模式^[1],与其它高性能计算技术相比,其基本特征是建立在大规模廉价服务器集群之上,通过基础设施与上层应用程序的协同构建达到最大效率利用硬件资源。我国城轨系统采用线网运营模式后,系统拓扑结构更为复杂,客流密度也更大,不同站点之间和不同线路之间相互影响,局部问题对整个线网运营的波及效应和联动效应更加突出,数据

处理规模也随之增加,加之各条线建设时间和设备厂家等问题导致数据的异构性使得系统集成性差,云计算可解决上述问题,利用大规模廉价服务器集群来协同完成数据的处理任务,信息分发是数据处理的重要部分^[2]。

1 相关工作

针对推拉式信息分发问题在学术界已经取得一定的研究成果。主要集中在 Web 环境下的信息分发,文献[3]提出根据发送数据块的往返时延决定网络中要传输的下一个节点的信息分发机制,主要用于视频流。文献[4]提出 CDN-P2P 和 Tree-Mesh 相结合的信息分发机制,也是用于 Web 中的视频流。文献[5]提出量子行为粒群优化算法的硬件实现,能够缩短信息的

收稿日期:2012-05-15;修回日期:2012-08-19

基金项目:甘肃省自然科学基金(ZS031-A25-019-G)

作者简介:韩泉叶(1974-),女,博士研究生,副教授,CCF 会员,主要研究领域为智能信息处理。

传送时间。文献[6]证明了在 N 个节点的传输网络中分发动态数据,树状网络结构的时间复杂性最低。文献^[7-9]也都研究的是 Web 环境下的视频流分发,文献[10]提出将系统中的异构、多层信息进行整合并为应用的集成提供了一种新思路。文献[11]论述了推拉技术在数字图书馆信息服务中的应用,文献[10,11]均没有从信息分发的及时性上进行分析。与这些相比,城轨线网的信息分发除了要从信息分发的及时性上进行分析外,还有其自身特点和结构。

其信息分发机制是关键。文中对云计算分布式编程模型中的推拉信息分发机制进行分析并将其应用在线网中,以达到信息分发时即便某个任务出现失败,也不会因数据的丢失而导致整个计算任务的重新执行^[12]。这一点能够很好地满足城轨线网大规模数据信息的完整性和准确性。

3 推拉信息分发机制

先看下面定义:

●暂时一致性需求 (Temporal Coherency Requirement, TCR): 定义用户或系统感兴趣的数据。其值 C 表示两次获取信息的最大可能偏差值, A_t 表示 Agent 在 t 时刻获取的信息值,保持暂时一致性也就意味着下列不等式一直存在。

$$||A_t - A_{t-1}|| \leq C \quad (1)$$

●刷新时间 (Time To Refresh, TTR): 表示 Agent 根据 TCR 计算的时间间隔,通过它,Agent 可以确定下一次从 Server 处获取数据信息的时间,然后刷新这期间变化的数据。

●及时性 (Currency): 公式(2)中, t_1, t_2, \dots, t_n 表示不满足用户一致性需求的时间间隔, T 表示用户观察的时间间隔。

$$\text{Currency} = 1 - \sum_{i=1}^n t_i / T \quad (2)$$

●通信量 (Communication overheads): 定义服务器和 Agent 之间交换数据信息的次数,推送时,网络中传输数据信息的次数等于推送的次数,而每一次拉取操作包括请求和应答。显然,对于确定数量的更新而言,拉取会产生更多的通信量。拉取时,Agent 根据当前数据变化率的快慢来估计何时获取 Server 数据信息值,如果数据变化率快,Agent 获取数据信息的时间间隔就会动态变小,结果就会导致网络中通信量增加。

●柔性 (Flexibility): 用来描述服务器出错情况下,推送与拉取的反应特性。推送时,如果服务器出错,Server 就不能主动地推送应该推送的信息,而用户却错误地认为系统没有可推送的数据信息;拉取时,如果出现服务器有问题,Server 被请求后没有应答,一段时间间隔后,系统就会知道服务器有问题而采取相应

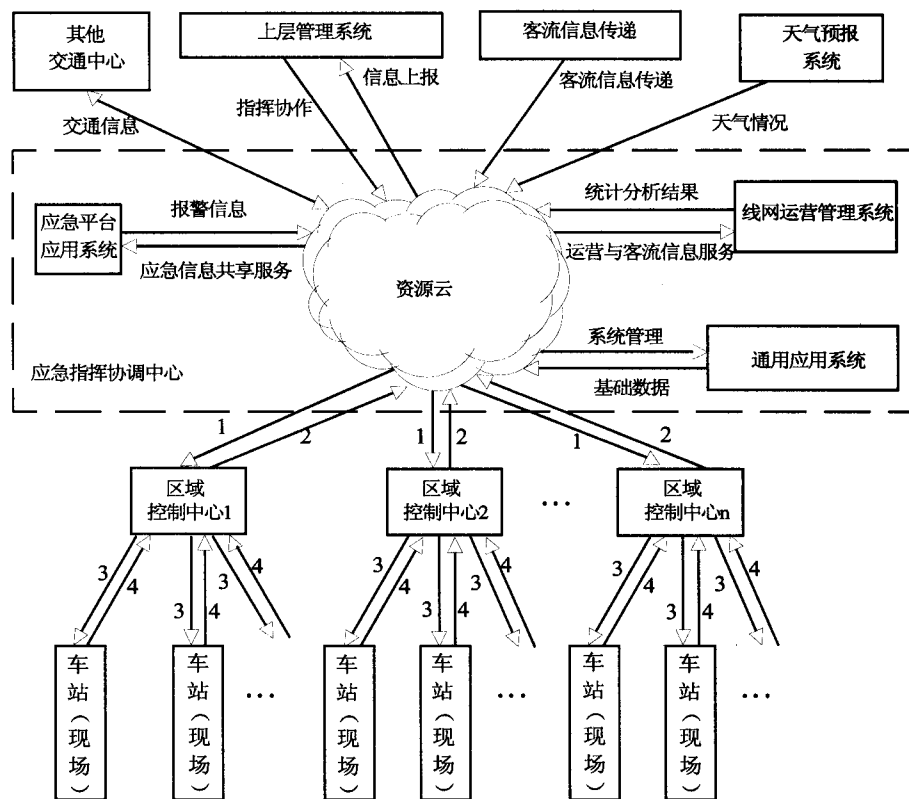


图1 城轨线网信息流图

2 城轨线网数据信息及其特征

城轨线网的信息流图如图1所示。其中应急指挥协调中心是线网运营的核心,承担着线网的信息共享、线网运营管理、应急处理和智能决策等重大任务;图中箭头表示信息流方向,1,2,3,4 分别代表四类信息流,其它类型的信息流已在图中注明,1 代表运营监督信息、决策指令信息、协调联动信息;2 代表监控信息、救援反馈信息、信息报送;3 代表资源调度信息、指令信息;4 代表监控信息、现场信息执行反馈。其中部分是实时信息流,即对信息的一致性要求十分严格,如报警信息、监控信息等;部分是非实时信息流,对信息的一致性要求不是十分严格,如乘客信息流、统计分析数据等;还有一些是根据需要随时获取的信息流,如公交线路图、半径覆盖范围内的 110、119 应急单位信息和现场视频流等。为了准确及时地处理大规模线网数据,

的措施。

3.1 拉取

拉取方法为了实现数据信息的暂时一致性, Agent 要根据数据信息变化率和用户一致性需求计算 TTR 值, 原则上一致性要求严格的数据信息, 对应一个较小的 TTR 值, 一致性要求不严格时, TTR 会变大。拉取是一种被动行为, 且 Agent 不需要拉取每一次变化的数据, 只需要拉取那些用户感兴趣且变化了的数据。因此拉取方法依赖于 TTR 值计算的准确性, 给定用户一致性需求, Agent 可以根据当前数据变化率动态计算 TTR 值, 变化率加快的数据会有一个动态变小的 TTR 值, 变化率变慢时, TTR 值会随之动态增加。TTR 的估计值 $TTR_{estimate}$ 可根据公式(3)计算。

$$\begin{cases} TTR_{estimate} = \frac{TTR_{last}}{|D_{last} - D_{nexttolast}|} \times C, \\ TTR_{min} \leq TTR_{last} \leq TTR_{max} \\ TTR_{last} = TTR_{estimate} \end{cases} \quad (3)$$

(3) 式中, TTR_{last} 表示上一次数据刷新的时间间隔, 其初始值为 $[TTR_{min}, TTR_{max}]$ 之间的一个随机值, $|D_{last} - D_{nexttolast}|$ 表示数据最后两次的差值。 $[TTR_{min}, TTR_{max}]$ 表示 TTR 的取值范围, C 表示数据信息 TCR 值。

3.2 推送

在推送方法中, 用户感兴趣的数据信息和有关的 TCR 值 C 需要注册在 Server 上, 无论何时, 只要用户感兴趣的数据信息发生了改变, 且满足式(4), Server 就会将此数据信息推送出去。

$$|D_p - D_{last}| \geq C, 0 \leq last < p \quad (4)$$

式(4)中, D_p 表示数据信息的当前值, D_{last} 表示 Server 最后推送的数据信息值。为了实现这个目标, Server 需要保存用户感兴趣数据的状态信息、该数据的 TCR 值和最后一次推送的该数据值。Server 能够主动识别数据的变化, 能够准确地确定何时应该推送什么样的数据, 因此推送方法适合于一致性需求非常严格的数据, 理论上只要 Server 正常, 推送数据的及时性就为 1, 推送是一种主动行为。

3.3 推送与拉取实验

实验中, 考虑到股市价格数据信息流与城轨线网数据信息流有相似处, 很多都属于实时信息流, 使用从 <http://www.kitco.com> 上获取的实时股票价格数据做实验, 以每秒 2 股的频率收集数据信息。拉取使用 HTTP Web Server 和代理作为实验平台, 推送使用单播和面向连接的 socket Server 做作为实验平台, 所有实验都在局域网上进行。拉取使用动态 TTR, TTR_{min} 等于 1 秒, 三个 TTR_{max} 值分别是 10、30 和 50 秒, C 等于 ¥0.06。

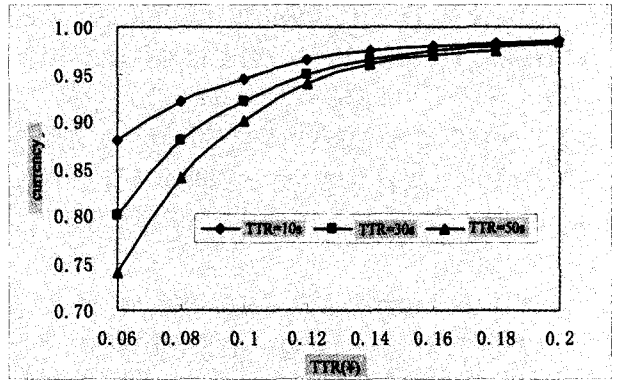


图2 拉取中 TCR 与及时性关系

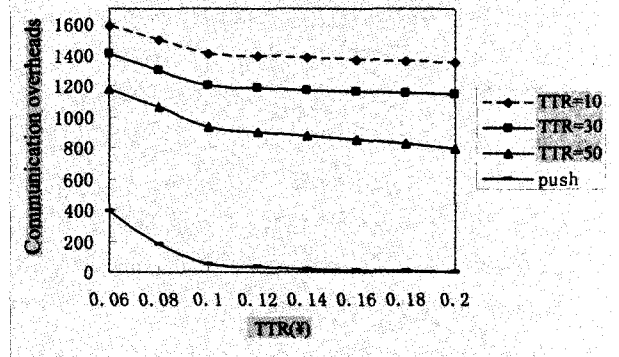


图3 TCR 与通信量关系

图2表明, 拉取给一致性需求严格的数据能提供 0.7~0.8 的及时性; 并且随着数据一致性要求的降低, 即随着 TCR 的增大, 及时性增大; 当 TCR 等于 ¥0.06 时, TTR 值小, 数据被刷新的频率快, 及时性值就大; 当 TCR 增大到 ¥0.2 时, 可以看到不同 TTR 得到几乎相同的及时性。图3表明, 推送产生的网络通信量较少, 原因是只有用户感兴趣的数据才被推送, 而拉取的网络通信量明显大于推送, 随着一致性需求的降低, 网络通信量也会降低。

从上述分析可知, 推送不但能提供高及时性还能降低网络通信量, 不失为一种好的信息分发方法, 但推送存在致命的弱点, 使用它会导致系统的柔性差, 使用推送方法, Server 需要保存客户的关键状态信息, 如用户的 TCR 信息, 若 Server 出错, 这种状态信息会丢失, 结果用户的 TCR 只有在 Server 被检测出出错并且重新注册 TCR 需求信息后, 才能重新被识别。这个问题可以在拉取方法中得到弥补, 若 Server 出错, 经过 TTR 后, Server 就会被检测出来。鉴于此种情况, 下面给出一种推拉结合的方法, 通过连接资源动态调节推送和拉取的信息分发机制。

4 推拉结合的信息分发机制

推拉结合的信息分发机制, 要求在任何时候, Server 要把它用户分成推送和拉取, 这种分类随着资源的变化而动态变化, Server 知道在某一时刻它能

控制的连接数目和能够提供给推送和拉取进行数据分发的资源数目, Server 给每一次连接提供的资源包括 sockets、memory、CPUtime 和 bandwidth。该机制的主要思想是根据资源情况和传输数据的及时性动态选择推送与拉取,从而提高系统柔性的同时减少数据块的分发延时。为了达到这一目标,当遇到下列情况之一时,需要使用推送。

(1) 检测当前数据值满足被推送的条件;

(2) 自上次推送后,经过一个特定的时间段 TTR。

前一个条件保证了推送数据的同步性,后一个条件保证了经过 TTR 后 Server 正常运行,增加了系统的柔性。当用户状态信息丢失或因网络问题导致无法连接 Server 时,经过 TTR,用户重新向 Server 请求恢复状态并拉取数据。这确保在最坏情况下,传输数据同步时间不会超过 TTR。根据资源情况动态调节推送与拉取,使用公式(5)、(6)调节。式(5)中, R_{present} 表示当前资源总数, N 表示用户总数, ξ 表示 pull 用户数量, S 表示资源 sockets 量, C 表示资源 CPUtime, M 表示资源 Memory, B 表示资源 bandwidth。

$$R_{\text{present}} = R_{\text{push}} + R_{\text{pull}}$$

$$= \left(\sum_{i=1}^{N-\xi} S_i, \sum_{i=1}^{N-\xi} C_i, \sum_{i=1}^{N-\xi} M_i, \sum_{i=1}^{N-\xi} B_i \right) + \left(\sum_{j=1}^{\xi} S_j, \sum_{j=1}^{\xi} C_j, \sum_{j=1}^{\xi} M_j, \sum_{j=1}^{\xi} B_j \right) \quad (5)$$

$$T_{\text{best}} = \min(T_{\text{best}}, \text{STD}_{\text{present}}) \quad (6)$$

式(6)中, STD (Sending Time Delay) 表示特定网络结构下数据块的发送时延, T_{best} 表示随着 ξ 变化最佳的 STD。推送和拉取之间的动态选择如图 4 所示:

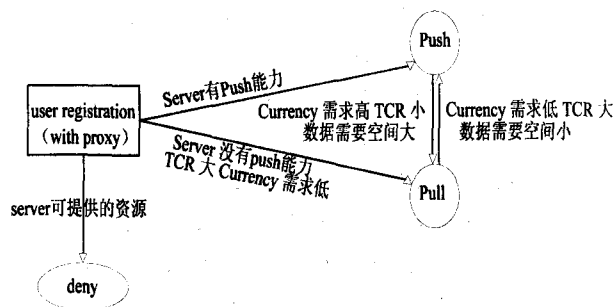


图 4 推送与拉取转化

图 4 中,不考虑数据信息的及时性,若 Server 有足够资源,则所有连接都初始化成推送;若用户能忍受的及时性低,那么此用户可使用拉取;若某用户需要 100% 的及时性,但 Server 没有足够资源,则 Server 将按步骤将一些推送转化为拉取,转化时需要考虑可提供的 bandwidth、数据变化率和 TCR。如果可提供的带宽较低,那么转化为拉取将会使现有情况变得更糟,因拉取比推送需要更多的带宽,如果存在及时性较低、数据变化率较低、TCR 较高、所需空间小的用户, Server

就会将其转化为拉取,当这种转变不可能时,该请求将被拒绝。

5 推拉结合的信息分发实验

实验中使用与上面相同的数据, Server 运行在 Linux 工作组平台上,用 4 个不同的工作组仿真客户机,每个客户机有 56 个用户,每次访问 3~4 个数据项,每次连接中,队列大小设置成 5,连接总数设置成 25。随着推送连接增加,由于资源受限导致请求被拒绝的情况如图 5 所示。图 5 中,当推送连接增加到 65% 以后,请求被拒绝的情况也随之增加,当 Server 接收到一个新的拉取请求时,如果队列里没有多余的空间,则请求不得不被拒绝,如果 65% 的连接是推送,意味着有 9 个连接是拉取,可容纳约 44 个拉取请求,当第 45 个拉取请求出现时就会被拒绝。当拉取请求被 Server 拒绝后,经过 TTR,被拒绝的用户会重新向 Server 发出请求,这会引起应答时间的增加,最终导致系统的及时性降低,如图 5 曲线的最后一部分所示。图 6 中,随着推送连接数的增加,系统的及时性也随之增加,当推送连接增加到 90% 附近时,由于系统资源受限导致系统的及时性下降。

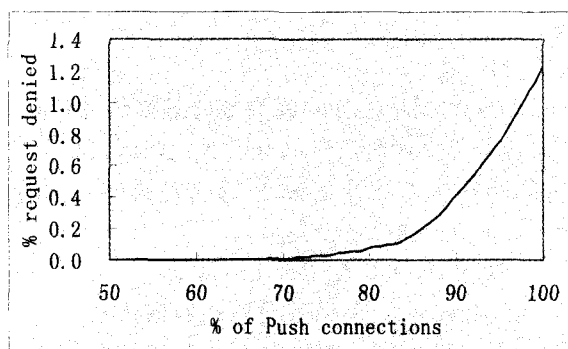


图 5 推送连接与请求拒绝关系

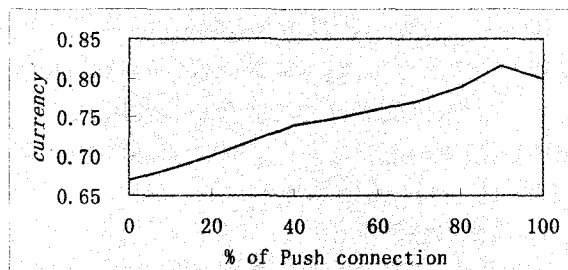


图 6 推送连接与系统及时性关系

6 结束语

对城轨线网数据信息及其特征进行了分析,讨论了云计算分布式编程模型中的推送和拉取信息分发机制,提出了一种适合于城轨线网的信息分发机制,该方法把推送与拉取相结合,通过动态分配可用资源来确

(下转第 210 页)

为 QPSK 时,在不同信噪比情况下,发送数据长度不同时,算法的平均误码率曲线如图 4 所示。

实验表明:发送数据长度不同时,采用改进激活函数的 CHNN 盲检测算法误码率大体相同,算法不会随着数据量的增大而失效。

4 结束语

CHNN 盲检测算法有效地利用了 Hopfield 神经网络的并行性、联想学习性等优点,较好地实现多值信号的盲检测,具有一定的应用价值。文中所提出的新的改进激活函数有效地提高了算法对噪声的抗干扰能力,适用于低信噪比、大数据量等复杂环境中,为进一步研究算法在物联网、传感网等复杂环境中的应用奠定了一定的基础。

参考文献:

- [1] Ding Z, Li Y. Blind Equalization and Identification [M]. New York: Marcel Dekker, 2002.
- [2] Giannakis G B, Hua Y B, Stoica P, et al. Signal Processing Advances in Wireless and Mobile Communications: Trends in Channel Estimation and Equalization [M]. NJ, USA: Prentice Hall PTR Upper Saddle River, 2000.
- [3] Santamaria I, Pantaleon C, Vielva L, et al. Blind equalization of constant modulus signal using support vector machines [J].

IEEE Transactions on Signal Processing, 2004, 52(6): 1773-1782.

- [4] 张志涌, 张 昀. 复数 Hopfield 盲恢复多用户 QPSK 信号 [J]. 东南大学学报, 2008, 38(Sup): 18-22.
- [5] 于舒娟, 张志涌. 含公零点 SIMO 信道 QPSK 序列盲检测 [J]. 东南大学学报, 2005, 35(6): 867-871.
- [6] Bai Erwei, Li Qingyu, Zhang Zhiyong. Blind source separation/channel equalization of nonlinear channels with binary inputs [J]. IEEE Transactions on Signal Processing, 2005, 53(7): 2315-2323.
- [7] Hopfield J J. Neurons with graded response have collective computational properties like those of two-state neurons [C]//Proceedings of the National Academy of Sciences of the United States of America. [s. l.]: [s. n.], 1984: 3088-3092.
- [8] Sengupta D, Iltis R A. Neural Solution to the Multitarget Tracking Data Association Problem [J]. IEEE Trans. on Aerospace and Electronic Systems, 1989, 25(1): 96-108.
- [9] Zaveri M, Merchant S N, Desai U B. Robust Neural-network Based Data Association and Multiple Model-based Tracking of Multiple Point Targets [J]. IEEE Transactions on Systems, 2007, 37(3): 337-351.
- [10] 敬忠良, 张国伟, 周宏仁. 随机神经网络在机动多目标跟踪中的应用 [J]. 航空学报, 1994, 15(7): 812-818.
- [11] Zurada J M. Neural Networks: Binary Monotonic and Multiple-valued [C]//Proc. of the 30th IEEE International Symposium on Multiple-valued Logic. Portland, Oregon: [s. n.], 2000.

(上接第 206 页)

定推送连接和拉取连接,最终达到在数据信息的及时性基础上,使数据块的发送时延达到最小,实验中的数据采用和城轨线网数据信息类似的股票数据信息,结果表明,推拉相结合的信息分发机制通过动态调节参数 ξ ,既可适用于城轨线网中一致性要求严格的数据信息也可适用于线网中一致性要求不太严格的数据信息,还可适用于线网中根据需要随时获取的数据信息。

参考文献:

- [1] 陈 康, 郑伟民. 云计算: 系统实例与研究现状 [J]. 软件学报, 2009, 20(5): 1337-1348.
- [2] Dean J, Ghemawat S. MapReduce: A Flexible Data Processing Tool [J]. Communications of the ACM, 2010, 53(1): 72-77.
- [3] Alessandro R, Renato L C. Delay-awareness in Push/Pull Streaming Protocols [R]. [s. l.]: [s. n.], 2010.
- [4] 黄思嘉, 吕智慧, 吴 杰. 新型双重混合的流媒体直播系统架构 [J]. 计算机工程, 2011, 37(9): 284-287.
- [5] 蔡 瑞, 须文波, 柴志雷, 等. 粒子群优化算法的硬件实现及其性能分析 [J]. 计算机工程, 2010, 36(4): 166-168.

- [6] Chakinala R. C, Kumarasubramanian A, Laing K A, et al. Playing Push vs Pull: Models and Algorithms for Disseminating Dynamic Data in Networks [C]//SPAA 2006. Cambridge, Massachusetts, USA: [s. n.], 2006.
- [7] Meng Z, Luo J G, Zhao L, et al. A Peer-to-Peer Network for Live Media Streaming-Using a Push-Pull Approach [C]//MM'05. Singapore: [s. n.], 2005.
- [8] 王 晖, 陈伟涛, 刘亚杰. 基于网络编码的 P2P 流媒体推拉结合数据调度方法 customR2 [J]. 计算机应用, 2010, 30(2): 285-288.
- [9] 卢良进, 万 健, 徐向华. 推拉结合的 P2P 直播系统研究与实现 [J]. 计算机工程, 2008, 34(8): 240-242.
- [10] 马忠贵, 叶 斌, 王成耀, 等. 基于“推-拉”技术的多源复杂信息整合研究 [J]. 计算机工程, 2006, 32(5): 178-180.
- [11] 李桂贞, 郑建明. 基于智能信息推拉技术的数字图书馆主动信息服务 [J]. 情报杂志, 2007(2): 65-67.
- [12] López-Neri E, Ramírez-Treviño A, López-Mellado E. Hierarchical Modeling of Urban Traffic Systems for Multi-Agent Based Simulation [C]//WCECS 2008. San Francisco, USA: [s. n.], 2008.