

完全服务和门限服务两级轮询系统 $E(x)$ 特性分析

杨志军^{1,2}, 丁洪伟², 陈传龙²

(1. 云南省教育厅教育科学研究院, 云南昆明 650223; 2. 云南大学信息学院, 云南昆明 650091)

摘要: 本文基于区分业务的优先级和提高系统的公平性, 提出了普通站点队列采用完全服务、中心站点采用门限服务策略的完全-门限服务两级轮询系统模型, 该模型既能满足区分业务优先级的服务需求, 又能兼顾系统的公平性。然后, 应用嵌入式马尔科夫链和概率母函数建立了系统的数学模型, 并精确解析了系统平均排队队长、平均查询周期等 $E(x)$ 特性。通过理论计算与仿真实验结果的对比, 说明了二者的一致性, 新的系统通过始终优先查询中心站点实现了区分优先级的控制, 同时普通站点采用完全服务, 在服务策略上得到较好的服务, 使系统有更好的公平性。

关键词: 轮询系统; 完全服务; 门限服务; 系统公平性

中图分类号: TP393

文献标识码: A

文章编号: 0372-2112(2014)04-0774-05

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2014.04.023

Research on $E(x)$ Characteristics of Two-Class Polling System of Exhaustive-Gated Service

YANG Zhi-jun^{1,2}, DING Hong-wei², CHEN Chuan-long²

(1. Educational and Scientific Institute, Educational Department of Yunnan Province, Kunming, Yunnan 650223, China;

2. School of Information Science and Engineering, Yunnan University, Kunming, Yunnan 650091, China;)

Abstract: Based on priority differentiation and fairness of the system, this paper proposes a two-class polling system, namely exhaustive-gated service system, which uses exhaustive service for ordinary stations and gated service for the central station. The scheme meets the requirements not only for priority differentiation but also fairness. The mathematical model of the new system is set up by method of imbedded Markov chain theory and generating function, then the $E(x)$ performance characteristics including mean queue length and polling cycle time are accurately analyzed. Theoretical and simulation results are identical and show that the new system efficiently differentiates priorities by always polling central station at first under better fairness for ordinary station.

Key words: polling system; exhaustive service; gated service; fairness

1 引言

轮询系统代表着一类调度控制模型, 因其公平性和对有限资源的合理分配与共享, 使其在现代计算机网络中得到广泛的应用, 网络媒介接入的控制方式通常分为轮询和随机多址两种主要方式。文献[1]综合阐述了轮询控制在计算机网络通信等各领域的应用, 文献[2~5]研究了其在无线局域网中的应用, 文献[6]研究了 Ad Hoc 网络中轮询调度策略的应用。轮询系统的基本模型由一个逻辑服务器和 N 个站点组成, 服务器按预定的服务规则依次查询各站点队列提供服务^[1]。轮询系统的性能通常由查询各站点的顺序、站点的服务策略和站点内信息分组的服务顺序等几个基本要素决定。按服

务策略轮询系统主要分为门限^[7]、完全^[8]和限定^[9]三类服务系统。研究人员围绕系统性能的优化与改进对各类轮询系统展开了研究, On Hashida 在文献[10]中应用嵌入式马尔可夫链和概率母函数分别建立了连续时间、对称的完全服务和门限服务系统模型, 文献[7~9]分别对门限服务、完全服务和限定服务系统的平均排队队长、平均等待时间等指标进行了深入分析, 文献[11]分析了无既定服务规则轮询系统的平均排队队长和负载。随着现代网络技术的快速发展, 单一的服务策略已不能够满足需要, 如多业务区分优先级的服务要求, 需要综合利用各种服务策略, 但这时系统分析的难度大大提高。文献[12]分析了混合服务方式的轮询系统, 但并未给出精确解析。文献[13]首次提出了“两级优先级控制轮询系

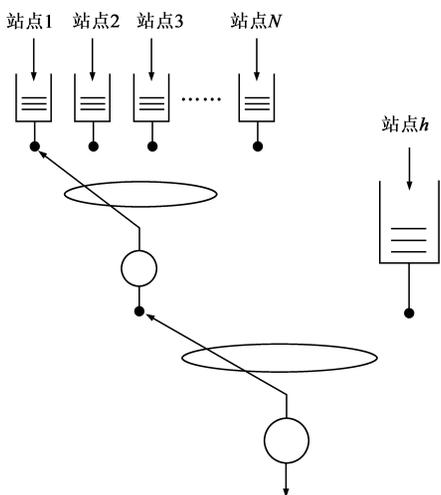


图1 两级优先级控制轮询系统模型

统模型”,通过中心站点采用完全服务、普通站点采用门限服务的两级轮询模式解决区分优先级的控制.接下来,文献[14]对离散时间两级完全服务轮询系统的理论进行了研究、文献[15]分析了离散时间的限定($K=1$)服务与完全服务的两级优先级轮询系统的控制模型、文献[16]分析了连续时间的限定($K=1$)服务与完全服务的两级轮询系统控制模型.以上两级轮询系统模型虽然通过两级轮询和不同服务策略解决了区分优先级控制的问题,但都是以中心站点采用完全服务的方式,降低了系统的公平性.

本文提出中心站点采用门限服务、普通站点采用完全服务的完全-门限服务两级轮询系统模型,既解决区分优先级服务问题,又兼顾了系统的公平性.并深入分析了系统的平均排队队长、查询周期和吞吐量等 $E(x)$ 特性.最后,通过仿真实验验证了系统的性能.

2 系统模型

2.1 模型原理

如图1所示,完全-门限服务两级控制轮询系统的原理为:轮询系统由 N 个普通站点队列和一个中心站点队列 h 两级组成,服务器对中心站点队列按门限服务规则进行服务,对普通站点队列按完全服务规则进行服务.服务器首先对中心站点队列服务,即只对开始查询服务时刻之前到达的信息服务,服务期间到达的则转入下一个轮询周期,若普通站点队列 i ($i=1,2,\dots,N$)不为空,则服务器按完全服务规则对 i 服务,当 i 的队列为空后,又转向查询中心站点队列 h ,中心站点队列完成规定的服务后,则开始对普通站点队列 $i+1$ 服务.完全-门限服务两级控制轮询系统,通过始终优先查询中心站点的服务区分了中心队列和普通队列,中心队列的服务得到优先保证,并通过普通站点采用

完全服务规则,而中心站点采用门限服务规则,兼顾了系统的公平性.

定义随机变量 $\xi_i(n)$ 是站点 i ($i=1,2,\dots,N$) 在 t_n 时刻其存储器内排队等待服务的信息分组数, $\xi_h(n)$ 表示中心站点 h 在 t_n 时刻其存储器内排队等待服务的信息分组数.整个系统在 t_n^* 时刻的状态变量为: $\{\xi_1(n), \xi_2(n), \dots, \xi_N(n), \xi_h(n)\}$, 在 t_n^* 时刻,系统的状态为: $\{\xi_1(n^*), \xi_2(n^*), \dots, \xi_N(n^*), \xi_h(n^*)\}$, 在 t_{n+1} 时刻整个系统的状态可表示为: $\{\xi_1(n+1), \xi_2(n+1), \dots, \xi_N(n+1), \xi_h(n+1)\}$.

可用马尔科夫链来描述系统的状态,该马尔科夫链是非周期的和各态历经的^[3,4,17].

2.2 变量定义

(1)每个站点信息分组的到达过程服从相互独立且同分布的泊松(Poisson)过程,在普通站点其分布的概率母函数、均值分别为 $A_i(z_i)$ 和 $\lambda_i = A_i'(z_i)$; 在中心站点 h 其分布的概率母函数、均值分别为 $A_h(z)$ 和 $\lambda_h = A_h'(1)$.

(2)服务器对任意站点一个信息分组的的服务时间服从相互独立、同分布的概率分布,在普通站点,其分布的概率母函数、均值分别为 $B_i(z_i)$ 和 $\beta_i = B_i'(1)$; 在中心站 h 其分布的概率母函数、均值分别为 $B_h(z_h)$ 和 $\beta_h = B_h'(1)$.

(3)任意普通站点完成相应的信息分组传输服务后,转向查询中心站点 h 的查询转移时间服从于一个相互独立、同分布的概率分布,其分布的概率母函数、均值分别为 $R(z)$, $\gamma = R'(1)$.

(4)普通站点对在任意时隙内到达的信息分组按完全服务策略服务所需的时间服从于一个相互独立、且同分布的概率分布,其分布的概率母函数为 $F(z)$.

定义以下变量:

u_i : 服务器从 i 号普通站点队列转向中心站点队列的查询转换时间;

v_i : 服务器对 i 号普通站点队列按完全服务的服务时间;

v_h : 服务器对中心站点队列按门限服务的服务时间;

$\mu_j(u_i)$: 在 u_i 时间内进入第 j 号站点队列 ($j=1,2,\dots,N,h$) 的信息分组数;

$\eta_j(v_i)$: 在 v_i 时间内进入第 j 号站点队列 ($j=1,2,\dots,N,h$) 的信息分组数;

$\eta_j(v_h)$: 在 v_h 时间内进入中心站点 h 的信息分组数.

根据模型原理,系统在各时刻的状态变量以下关

系:

$$\begin{cases} \xi_{ih}(n) = \xi_h(n) + \mu_h(u_i) + \eta_h(v_i) \\ \xi_k(n+1) = \xi_k(n) + \mu_k(u_i) + \eta_k(v_h) & k=1,2,\dots,N, k \neq i \\ \xi_i(n+1) = \eta_i(v_h) + \mu_i(u_i) \end{cases} \quad (1)$$

2.3 概率母函数

假定每个普通站点和中心站点的存储器容量足够大,不会产生信息分组丢失,并且信息分组的服务按先到先服务(FCFS)的原则服务,在 $\sum_{i=1}^N \lambda_i \beta_i + \lambda_h \beta_h < 1$ 的条件下系统达到稳态^[3,17]. 稳态下的概率母函数定义为:

$$\begin{aligned} \lim_{n \rightarrow \infty} P[\xi_j(n) = x_j, j=1,2,\dots,N, h] &= \pi_i(x_1, x_2, \dots, x_i, \dots, x_N, x_h) \\ G_i(z_1, z_2, \dots, z_N, z_h) &= \sum_{x_1=0}^{\infty} \sum_{x_2=0}^{\infty} \dots \sum_{x_i=0}^{\infty} \dots \sum_{x_N=0}^{\infty} \pi_i(x_1, x_2, \dots, x_i, \dots, x_N, x_h) \\ &\quad \cdot z_1^{x_1} z_2^{x_2} \dots z_i^{x_i} \dots z_N^{x_N} z_h^{x_h} \\ &\quad i=1,2,\dots,N \end{aligned} \quad (2)$$

根据式(1)和式(2),两级轮询系统在 t_{n^*} 时刻服务中心站点时的概率母函数为:

$$\begin{aligned} G_{ih}(z_1, z_2, \dots, z_N, z_h) &= \lim_{t \rightarrow \infty} E \left[\prod_{j=1}^N z_j^{\xi_j(n^*)} z_h^{\xi_h(n^*)} \right] \\ &= R \left(\prod_{j=1}^N A_j(z_j) A_h(z_h) \right) G_i[z_1, z_2, \dots, z_{i-1}, \\ &\quad B_i \left(\prod_{j=1}^N A_j(z_j) A_h(z_h) F_i \left(\prod_{j=1}^N A_j(z_j) A_h(z_h) \right) \right), \\ &\quad z_{i+1}, \dots, z_N, z_h] \\ &\quad i=1,2,\dots,N \end{aligned} \quad (3)$$

系统在 t_{n+1} 时刻服务 $i+1$ 号普通站点时的概率母函数为:

$$\begin{aligned} G_{i+1}(z_1, z_2, \dots, z_n, z_h) \\ &= \lim_{t \rightarrow \infty} E \left[\prod_{j=1}^N z_j^{\xi_j(n+1)} z_h^{\xi_h(n+1)} \right] \\ &= G_{ih}[z_1, z_2, \dots, z_N, B_h \left(\prod_{j=1}^N A_j(z_j) A_h(z_h) \right)] \\ &\quad i=1,2,\dots,N \end{aligned} \quad (4)$$

3 模型分析

3.1 系统平均排队队长

定义在 t_n 时刻第 i 号节点开始传输服务时,第 j 号节点存储器中平均存储的信息分组数为 $g_i(j)$,则:

$$\begin{aligned} g_i(j) &= \lim_{z_1, \dots, z_N, z_h \rightarrow 1} \frac{\partial G_i(z_1, z_2, \dots, z_N, z_h)}{\partial z_j} \\ &\quad i=1,2,\dots,N, h; j=1,2,\dots,N, h \end{aligned} \quad (5)$$

由式(3)、(4)计算得到

普通站点的平均排队队长为:

$$g_i(i) = \frac{N\gamma\lambda - N\rho\gamma\lambda}{1 - \rho_h - N\rho} \quad (6)$$

中心站点的平均排队队长为:

$$g_{ih}(h) = \frac{\lambda_h\gamma}{1 - N\rho - \rho_h} \quad (7)$$

其中 $\rho_h = \lambda_h\beta_h, \rho = \lambda\beta$.

3.2 系统平均查询周期

系统平均查询周期指两次查询同一站点间的平均时间.

普通站点的平均查询周期为:

$$E(\theta_i) = \frac{N\gamma}{1 - N\rho - \rho_h} \quad (8)$$

中心站点查询周期:

$$E(\theta_h) = \frac{\gamma}{1 - N\rho - \rho_h} \quad (9)$$

3.3 系统吞吐量

定义:系统吞吐量 $E(T) = \frac{\text{总服务时间}}{\text{系统运行总时间}}$

$$E(T) = N\rho + \rho_h \quad (10)$$

4 仿真实验及 $E(x)$ 特性分析

根据以上所建立的完全-门限服务两级轮询系统模型,在系统稳定条件下 $\sum_{i=1}^N \lambda_i \beta_i + \lambda_h \beta_h < 1$ 进行数值计算和仿真实验.

从图2至图9中可以得出理论计算与仿真实验的结果是一致的,通过实验验证了系统模型的合理性,并可进一步分析系统的性能:

(1)图2与图3描述了系统的平均排队队长、平均查询周期与系统负载的关系,可以看到,中心站点与普通站点的平均排队队长、平均查询周期得到了很好的区分,在系统负载增加的情况下,中心站点的平均排队队长、平均查询周期仍然比普通站点的小,得到了明显区分.这进一步说明,虽然普通站点采用了平均等待时延更小的完全服务,但通过访问频率的保证使得中心站点依然享有更高的服务优先级.

(2)图4与图5给出了系统的平均排队队长、平均查询周期与服务时间的关系,可以看到信息分组的平均排队队长、平均查询周期随服务时间变化的趋势,与上面类似,中心站点与普通站点得到了明显的区分,同样在服务时间增加的情况下,也能保证中心站点的平均排队队长、平均查询周期比普通站点的小.

(3)图6与图7分别反映了系统吞吐量与到达率和服务时间的关系,再次说明理论计算值与实验值的一致性,同时可以看到吞吐量随着到达率、服务时间增加的合理增加.

(4)图8与图9分别展示了门限-完全服务两级轮询系统(即中心站点采用完全服务策略,而普通站点采用门限服务策略的系统)与本文描述的完全-门限服务两级轮询系统的比较.图8反映出门限-完全服务系统中心站点的平均排队队长要低于完全-门限服务系

统,但图 9 反映出门限-完全服务系统普通站点的平均排队队长要高于完全-门限服务系统,结合图 2 中心站点与普通站点的区分,充分说明本文描述的完全-门

限服务系统在优先满足了中心站点服务的同时,又兼顾了普通站点的服务,具有更好的公平性。

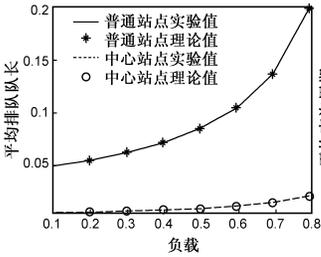


图2 平均排队队长与系统负载的关系

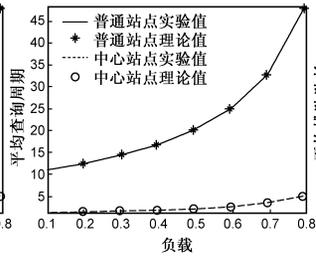


图3 平均查询周期与系统负载的关系

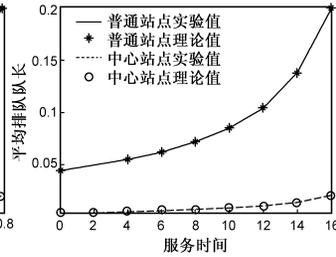


图4 平均排队队长与服务时间的关系

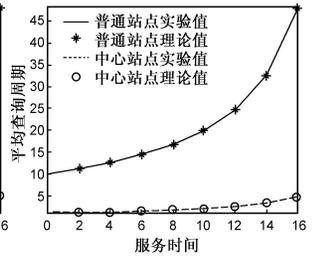


图5 平均查询周期与服务时间的关系

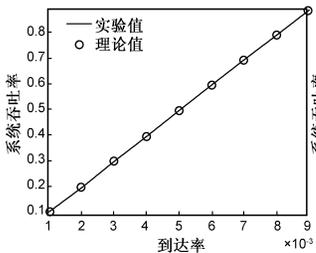


图6 系统吞吐量与到达率的关系

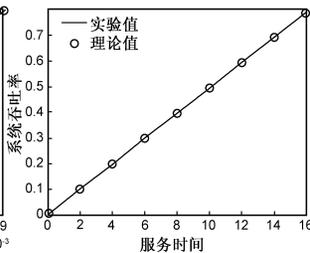


图7 系统吞吐量与服务时间的关系

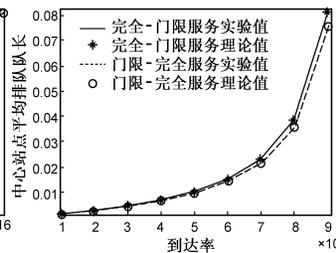


图8 二类两级轮询系统中心站点比较

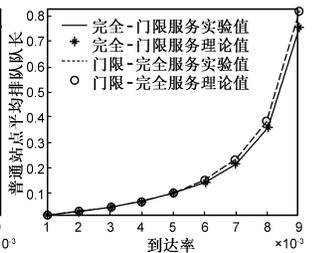


图9 二类两级轮询系统普通站点比较

5 结论

在解决区分优先级控制的问题时,需要同时考虑系统的公平性,以使系统具有更好的效率与合理性,以满足实际应用的需求。

本文提出了普通站点队列采用完全服务,中心站点队列采用门限策略的完全-门限服务两级优先级控制轮询系统模型,既实现了区分优先级的控制,又兼顾了普通站点的服务.该系统通过始终查询中心站点保障了中心站点的优先级服务,然而,又为普通站点提供时延较小的完全服务,在保证优先级服务的同时体现了更好的公平性.采用嵌入式马尔可夫链和概率母函数的方法建立了系统模型,精确计算了系统的平均排队队长、平均查询周期和吞吐量等 $E(x)$ 特性,通过仿真实验验证了理论模型.理论计算和实验结果充分说明中心站点和普通站点性能得到了很好的区分,保证了中心站点队列服务的优先级,而普通站点采用完全服务策略也得到了较好的服务,与以往两级轮询系统不同的是,本系统在同时满足中心站点高优先级应用需求的同时,又能兼顾了普通站点的服务,使系统具有更好的公平性,拓展了两级优先级控制轮询系统的研究和应用空间.今后的研究工作将进一步考虑中心站点设置的动态性和服务策略的灵活性。

参考文献

- [1] M A A Boon, R D van der Mei, E M M Winands. Applications of polling systems [J]. *Surveys in Operations Research and Management Science*, 2011, 16(2): 67 - 82.
- [2] Eustathia Ziouva, Theodore Antonakopoulos. A dynamically adaptable polling scheme for voice support in IEEE 802.11 networks [J]. *Computer Communications*, 2003, 26: 129 - 142.
- [3] Yang Zhijun, Zhao Dongfeng. Polling strategy for wireless multimedia LANs [J]. *Tsinghua Science and Technology*, 2006, 11(5): 606 - 610.
- [4] 张宇眉, 赵东风. 无线局域网中动态优先级轮询控制系统分析 [J]. *计算机工程与应用*, 2005, 03: 122 - 124.
Zhang Yumei, Zhao Dongfeng. Analysis of dynamic priority polling system based on wireless LANs [J]. *Computer Engineering and Applications*, 2005, 03: 122 - 124. (in Chinese)
- [5] Yang Zhijun, Zhao Dongfeng. A new priority-based scheme for QoS differentiation in wireless LAN [A]. *Proceedings of Workshop on Intelligent Information Technology Application, IITA 2007 [C]*. China: IEEE Computer Society, 2007. 192 - 195.
- [6] 何敏, 刘心松, 赵东风, 刘克剑, 陈建英. 移动 ad hoc 网络轮询接入控制协议 [J]. *电子与信息学报*, 2005, 27(7): 1147 - 1152.
He Min, Liu Xin-song, Zhao Dong-feng, Liu Ke-jian, Chen Jian-ying. A polling access control protocol in mobile ad hoc networks [J]. *Journal of Electronics and Information Technolo-*

- gy, 2005, 27(7): 1147 - 1152. (in Chinese)
- [7] 赵东风, 郑苏民. 周期查询式门限服务排队系统中信息分组的延迟分析[J]. 通信学报, 1994, 15(2): 18 - 23.
Zhao Dongfeng, Zheng Sumin. Message waiting time analysis for a polling system with gated service[J]. Journal of Communications, 1994, 15(2): 18 - 23. (in Chinese)
- [8] 赵东风, 郑苏民. 查询式完全服务排队模型分析[J]. 电子学报, 1994, 22(5): 102 - 107.
Zhao Dongfeng, Zheng Sumin. Analysis of a polling model with exhaustive service[J]. Acta Electronica Sinica, 1994, 22(5): 102 - 107. (in Chinese)
- [9] 赵东风, 李必海, 郑苏民. 周期查询式限定服务排队系统研究[J]. 电子科学学刊, 1997, 19(1): 44 - 49.
Zhao Dongfeng, Li Bihai, Zheng Sumin. Study of polling systems with limited service[J]. Journal of Electronics, 1997, 19(1): 44 - 49. (in Chinese)
- [10] O Hashida. Analysis of multiqueue[J]. Review of the Electrical Communication Laboratories, 1972, 20(3 - 4): 189 - 199.
- [11] O J Boxma, O Kella, K M Kosinski. Queue lengths and workloads in polling systems[J]. Operations Research Letters, 2011, 39(6): 401 - 405.
- [12] Hwang L C. An exact analysis of an asymmetric polling system with mixed service discipline and general service order[J]. Computer Communication, 1997, 20(10): 1293 - 1299.
- [13] 杨志军, 赵东风, 丁洪伟, 赵一帆. 两级优先级控制轮询系统研究[J]. 电子学报, 2009, 37(7): 1452 - 1456.
Yang Zhijun, Zhao Dongfeng, Ding Hongwei, Zhao Yifan. Research on two-class priority based polling system[J]. Acta Electronica Sinica, 2009, 37(7): 1452 - 1456. (in Chinese)
- [14] 梁竹关. 离散时间完全服务两级轮询系统理论研究[D]. 昆明: 云南大学, 2010.
Liang Zhuguan. Research on the theory of discrete-time two-class polling system with exhaustive service[D]. Kunming: Yunnan University, 2010. (in Chinese)
- [15] 柳虔林. 离散时间完全与限定(K=1)服务两级轮询系统理论研究[D]. 昆明: 云南大学, 2010.

Liu Qianlin. Researches on discrete-time two-level-polling system with exhaustive and limited(K=1) service[D]. Kunming: Yunnan University, 2010. (in Chinese)

- [16] 保利勇. 连续时间的优先级完全与限定服务轮询系统研究[D]. 昆明: 云南大学, 2011.

Bao Liyong. Researches on continuous-time priority polling system under the exhaustive and limited service policy[D]. Kunming: Yunnan University, 2011. (in Chinese)

- [17] O C Ibe, Xian Cheng. Stability conditions for multiqueue systems with cyclic service[J]. IEEE Tran Automat Control, 1988, 33(1): 102 - 103.

作者简介



杨志军 男, 1968年5月出生于云南省保山市, 研究员, 云南省教育厅教育科学研究院副院长, 云南大学信息学院硕士生导师. 1990年毕业于浙江大学计算机科学与工程系, 获工学学士学位, 2002年于云南大学信息学院获通信与信息系统专业工学硕士学位, 2008年在云南大学信息学院获通信与信息系统专业工学博士学位. 研究方向: 计算机通信与网络、无线通信、轮询系

统和教育信息化.

E-mail: yzj@ynjy.cn



丁洪伟 男, 1964年6月生于云南省景洪市, 云南大学信息学院副教授、博士, 主要研究方向为轮询系统、随机多址通信系统.

E-mail: dhw1964@163.com



陈传龙 男, 1985年11月生于安徽省怀宁县, 云南大学信息学院通信与信息系统专业硕士研究生, 主要从事无线通信网络、轮询多址理论的研究.

E-mail: aqcl@126.com