

交换式以太网实时消息可调度性分析

檀明^{1,2}, 魏臻¹, 韩江洪¹, 钟伯成²

(1 合肥工业大学 计算机与信息学院, 安徽 合肥 230009; 2 合肥学院 网络与智能信息处理重点实验室, 安徽 合肥 230601)

摘要: 针对交换式以太网不能满足工业数据通信的实时性要求问题, 首先提出了一种改进的, 在站点和交换机端均使用 EDF 算法的消息调度模型, 并由源站点和目标站点以分布式方式对实时消息在链路的可调度性进行判定, 从而简化了交换机的设计. 另外, 为保证实时消息在交换式以太网传输的实时性, 分别给出了实时消息在发送和接收链路的可调度性判定条件, 并进行了相应的理论证明.

关键词: 交换式以太网; EDF 调度算法; 周期性实时消息

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2011)03-0051-04

Feasibility Analysis for Real-Time Messages over Switched Ethernet

TAN Ming^{1,2}, WEI Zhen¹, HAN Jiang-hong¹, ZHONG Bo-cheng²

(1 School of Computer and Information, HFUT, Hefei 230009, China;

2 Key Laboratory of Network and Intelligent Information Processing, Hefei University, Hefei 230601, China)

Abstract: To overcome the drawback that switched Ethernet can not meet the real-time requirements in industrial communications, an improved message transmission model in which switch and the end-nodes controls the real-time traffic with Earliest Deadline First (EDF) scheduling was proposed. In addition, a distributed feasibility analysis method which is performed by the source and the destination nodes without the modification of the operational features of Ethernet switch was presented. Moreover, a more general condition for feasibility analysis for real-time periodic message over switched Ethernet was given and proved by using real-time theory.

Key words: switched Ethernet; EDF-based scheduling algorithm; real-time periodic message

1 引言

为了能将交换式以太网用于工业实时通信领域, 文献[1]将网络演算理论(Network Calculus)用于交换式以太网实时性能评价. 通过分析数据流在服务站点的到达曲线和服务曲线, 计算其服务时延, 得到了不同优先级数据流的最大网络时延, 但由于总是假定各个站点具有同时突发的数据, 得到的最大时延过于保守. 因此文献[2]提出传输整形方法, 通过在每个站点的传输层(如 TCP/UDP)与 Ethernet 之间增加了一个流量整形器, 以使网络的平均负载始终保持在某个较低的规定值, 利用低负载、低冲突概率原理来保证消息传输时间的可预测性, 但

无法保证严格的实时性传输要求.

文献[3]针对交换式全双工以太网, 通过在交换机和站点的网络层与链路层之间增加通信控制层(即 RT 层), 改进交换机硬件以支持实时数据传输, 由交换机负责实时通道的建立, 站点通过实时通道(任两个站点间存在的虚拟链路)相互通信. 交换机和站点均采用 EDF(Earliest Deadline First)算法调度实时消息的传输, 较好地解决了网络消息的实时性传输问题, 但仍存在以下不足, 有些假定和条件往往与实际网络不符, 另外文献[3]推导出的可调度性判定方法缺乏理论依据.

本文改进了文献[3]提出的实时消息调度模型, 将对实时消息的可调度性分析由交换机端的集中判

定改为由源站点和目标站点以分布方式协同完成,简化了交换机的设计. 针对文献[3]在实时消息调度的假设条件方面的不足进行了修正:假定实时消息的截止期小于或等于其周期(这种假设在实际应用中一般是成立的),站点的实时消息调度采取不可抢占的 EDF 算法,对于接收链路的可调度性分析,不要求实时消息满足周期性释放条件. 在此基础上,本文重新研究了实时消息在交换式以太网上的可调度判定条件并给出了相应证明.

2 网络结构

采取全双工、微网段的通信方式,如图 1 所示. 每个站点仅与一个交换机的端口相连且具有一个发送链路(TL)和一个接收链路(RL),每条链路可同时进行以相同速率进行数据的发送和接收. 交换机和终站点均带有实时通信层(RT 层),站点通过实时通道(站点间存在的虚拟链路)相互通信. 实时消息使用 UDP 协议传输,非实时消息使用 TCP 协议传输. 交换机和站点的 RT 层均设有两个消息缓冲队列,其中实时消息队列按 EDF 顺序传输,非实时消息队列按 FCFS(First Come First Served)顺序传输,实时消息队列的传输优先级高于非实时消息队列,只有实时消息队列为空时才能开始非实时消息队列的传输(如图 2 所示).

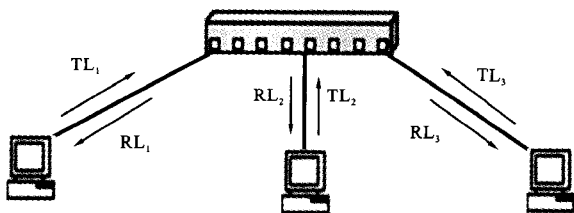


图 1 网络通信模型

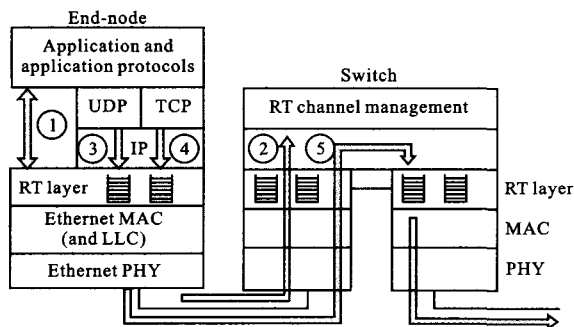


图 2 RT 层和输出队列示意图

3 实时调度模型

交换机以 T_{SYN} 为间隔,周期性地向所有站点发

送时钟同步消息(SYN),从而使网络站点以 T_{SYN} 为精度实现时钟同步. 由于在典型的工业实时通信中传输的主要为周期性消息(非周期性的突发实时消息因其一般存在最小的产生周期,从而可转化为周期性实时消息加以处理),而非实时消息(NRT)则无传输时限要求,因而本文只研究周期性实时消息(PRT)的调度问题. 周期性实时消息集合表示为

$$\text{PRT} = \{M_i, i = 1, 2, \dots, n\}, M_i = (D_i, T_i, O_i, C_i),$$

其中, D_i, T_i, O_i, C_i 分别为实时消息 M_i 的相对截止期、产生周期、初始相位和传输时间, n 为周期性实时消息的数量. 假定 D_i, T_i, O_i 均为时钟同步周期 T_{SYN} 的整数倍,且 $D_i \leq T_i$. M_i 从源站点产生到被目标站点接收的总时延 $T_{\text{delay}} = W_{\text{src}} + W_{\text{dest}} + 2 * t_p + 2 * C_i + t_s$,其中 $W_{\text{src}}, W_{\text{dest}}$ 分别对应实时消息 M_i 在源站点和交换机端等待被传输的时间, t_p 为链路上消息的传播时延; C_i 为消息传输时延,取决于消息的长度与链路带宽; t_s 为交换机基本转发时间,包含目标端口的查找时间和交换单元的建立时间. 因 $t_p + t_s$ 一般可视为常量,不失一般性,故本文在分析时予以忽略.

当源站点需要向目标站点发送实时消息 M_i 时,首先向本地 RT 层发出请求,由本地 RT 层对 M_i 进行发送链路可调度性分析,当 M_i 在发送链路可调度时, M_i 的发送请求通过交换机将被继续向目标站点的 RT 层转发,由目标站点 RT 层继续对 M_i 在接收链路进行可调度性分析,当 M_i 在发送和接收链路均可调度时,目标站点 RT 层通知源站点接纳实时消息 M_i ,否则拒绝 M_i 的发送请求.

与文献[3]不同,本文假定源站点的 RT 层对于周期性实时消息采取同步调度模型^[4],消息被分配到一系列调度窗口中进行传输,调度窗口大小设置为 T_{SYN} (如图 3 所示). 在每个 T_{SYN} 的开始(当前 SYN 到达时刻),调度器对实时消息按绝对截止期由小至大进行排序,形成 EDF 调度队列. 依次对 EDF 调度队列中的实时消息进行边界检查,只有满足在当前调度窗口(下一 SYN 到达之前)内能在发送链路上完成传输的实时消息才会被调度. 当没有实时消息可供调度时,调度器将从非实时消息缓冲队列中选择能在本调度窗口完成发送的消息进行传输. 消息到达交换机的对应输出端口缓冲队列后,由交换机负责调度消息在接收链路上进行传输. 其中时间同步消息 SYN 的优先级最高,可以抢占实时消息与非实时消息的传输(如图 4 所示). 只有当 SYN 在接收链路完成传输后,

实时消息和非实时消息才能被调度传输. 同样, 实时消息的优先级高于非实时消息, 可以抢占非实时消息的传输(如图5所示). 但是, 实时消息集内可能会出现高优先级消息被低优先级消息阻塞的情况, 即当某一高优先级实时消息到达交换机端时, 若接收链路恰好被低优先级的实时消息所占用, 只能等低优先级的实时消息传输完成后, 高优先级消息才能按 EDF 算法被调度传输.

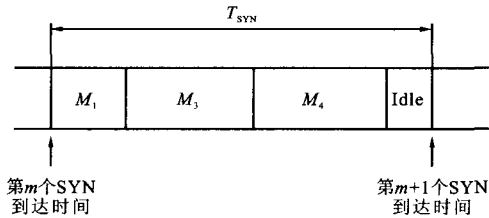


图3 消息在发送链路上的传输

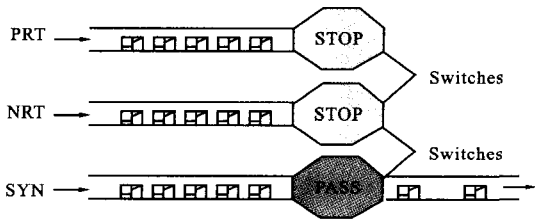


图4 SYN消息可抢占实时和非实时消息

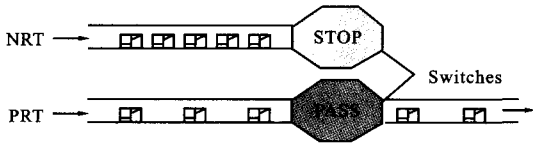


图5 实时消息可抢占非实时消息

4 可调度性分析

周期性实时消息 M_i 从源站点到目标站点的传输过程, 可分解为从源站点到交换机的发送阶段和从交换机到目标站点的接收阶段, 从而对实时消息 M_i 的可调度分析也可分解为对发送链路实时性要求 $M_{i,1}(T_i, C_i, D_{i,1}, O_i)$ 的可调度分析和对接收链路实时性要求 $M_{i,2}(T_i, C_i, D_{i,2}, O_i + D_{i,1})$ 的可调度分析两部分, 其中 $D_{i,1}$ 为实时消息 M_i 在发送阶段的相对截止期, $D_{i,2}$ 为实时消息 M_i 在接收阶段的相对截止期, 且满足 $D_i = D_{i,1} + D_{i,2}$. 假定实时消息 M_i 在接收链路调度采取非连续型同步模式, 即实时消息 M_i 第 j 个实例对应接收链路的释放时间 $R_{i,j} = j * T_i + D_{i,1}$, 从而消息在接收链路的调度满足周期性释放条件. 显然, 如果实时消息 M_i 均能在 $D_{i,1}$ 和 $D_{i,2}$ 内分别在发送和接收链路被调度传输, 则 M_i 实时性可以得到保证.

引理1 单处理器上的周期性实时任务集, 如按照 EDF 调度策略可调度, 则在绝对截止期不变的前提下, 作业的提前释放不影响其可调度性^[5].

根据本文调度模型, 实时消息 M_i 在接收链路的释放时间为发送链路的传输完成时间, 由于在发送链路调度所经时间的不同, 其在接收链路上调度不再满足周期性释放条件, 但实时消息 M_i 的绝对截止期的出现满足周期性. 将对实时消息 $M_i(T_i, D_i, C_i, O_i)$ 的调度可分解成: $M_{i,1}(T_i, C_i, D_{i,1}, O_i)$ 和 $M_{i,2}(T_i, C_i, D_{i,2}, O_i + D_{i,1})$, 其中 $D_i = D_{i,1} + D_{i,2}$, $D_{i,1} = \lfloor D_i/2 \rfloor$, $D_{i,2} = \lceil D_i/2 \rceil$. 当限定实时消息 M_i 第 j 个实例对应接收链路的释放时间 $R_{i,j} \leq j * T_i + D_{i,1}$, 根据引理1, 如果 $M_{i,1}$ 和 $M_{i,2}$ 分别在发送和接收链路满足可调度条件, 则采取连续型同步模式(对消息 $M_{i,2}$ 的释放时间不做假设), 实时消息 M_i 也可被调度.

4.1 消息在发送链路可调度性分析

对于相对截止期小于或等于其周期的实时任务集调度问题, 文献[6]中给出了任务集在单处理器上基于可抢占式 EDF 调度算法可调度性判定的充分条件(如定理1所示).

定理1 $TS = \{TS_i, i = 1, 2, \dots, n\}$ 是包含 n 个任务的任务集, 其中 $TS_i = (D_i, T_i, C_i)$, $i = 1, 2, \dots, n$ 是一系列按相对截止期非递减顺序排列的周期性实时任务, C_i 为任务 TS_i 的最坏执行时间, T_i 为任务 TS_i 的周期, D_i 为 TS_i 的相对截止期, 且 $D_i \leq T_i$, 当满足式(1)时, 基于抢占式 EDF 调度算法 TS 可调度.

$$\forall (1 \leq k \leq n) \left| \left(\sum_{i=1}^k \frac{C_i}{T_i} + \frac{1}{D_k} \left(\frac{T_i - \min(T_i, D_i)}{T_i} \right) \cdot C_i \leq 1 \right) \right. \quad (1)$$

定理2 $PRT = \{M_i, i = 1, 2, \dots, n\}$ 是包含 n 个实时消息的消息集, 其中 $M_i = (D_i, T_i, C_i)$, $i = 1, 2, \dots, n$ 是一系列按相对截止期非递减顺序排列的周期性实时消息, D_i, T_i, C_i 分别为实时消息 M_i 的相对截止期、产生周期和传输时间, 且 $D_i \leq T_i$, 当满足式(2)时, 对于本文发送链路调度模型, 消息集 PRT 可被调度.

$$\forall (0 \leq k \leq n) \left| \left(\sum_{i=0}^k \frac{C_i}{T_i} + \frac{1}{D_k} \left(\frac{T_i - \min(T_i, D_i)}{T_i} \right) \cdot C_i \leq 1 \right) \right. \quad (2)$$

式中, $D_0 = T_0 = T_{\text{SYN}}, C_0 = \max\{C_i\}$.

证明: 根据源站点 RT 层的发送链路调度模型, 虽然消息在发送链路的传输不可抢占, 但由于调度

窗口内部不会有实时消息产生(实时消息均在调度窗口左边界到达并等待被调度),且被调度的消息在发送链路的传输完成时间不会跨越调度窗口的右边界,在调度窗口内部不会发生低优先级消息被更高优先级消息抢占的情况,因而对 PRT 可调度分析可按基于可抢占式 EDF 调度算法进行。

为保证被调度的消息在发送链路的传输完成时间不会跨越调度窗口的右边界,最坏情况是每个调度窗口均要插入长度为 $\max\{C_i\}$ 的传输空隙(如图 3 所示)。为此,引入虚拟周期性实时消息 $M_0 = (D_0, T_0, C_0)$ 表示这些传输空隙,其中 $D_0 = T_0 = T_{\text{SYN}}$, $C_0 = \max\{C_i\}$ 。显然虚拟消息 M_0 具有最高优先级。为使 $\text{PRT} = \{M_i, i = 1, 2, \dots, n\}$ 可调度,只要使 $\text{PRT}_0 = \{M_i, i = 0, 1, \dots, n\}$ 可调度。由定理 1 得到其可调度的充分条件如式(2)所示。

4.2 消息在接收链路可调度性分析

对于相对截止期小于或等于其周期的实时任务集调度问题,文献[4]中给出了任务集在单处理器上基于非可抢占式 EDF 调度算法可调度性判定的充分条件(如定理 3 所示)。

定理 3 $\text{TS} = \{TS_i, i = 1, 2, \dots, n\}$ 是包含 n 个任务的任务集,其中 $TS_i = (D_i, T_i, C_i), i = 1, 2, \dots, n$ 是一系列按相对截止期非递减顺序排列的周期性实时任务, C_i 为任务 TS_i 的最坏执行时间, T_i 为任务 TS_i 的周期, D_i 分别为 TS_i 的相对截止期,且 $D_i \leq T_i$, 当满足式(3)时,基于非抢占式 EDF 调度算法 TS 可调度。

$$\forall (0 \leq k \leq n) \left| \left(\sum_{i=0}^k \frac{C_i}{D_i} + \frac{B_k}{D_k} \leq 1 \right) \right. \quad (3)$$

式中, B_k 为任务 TS_k 被低优先级任务阻塞的最大时间。

定理 4 $\text{PRT} = \{M_i, i = 1, 2, \dots, n\}$ 是包含 n 个实时消息的消息集,其中 $M_i = (D_i, T_i, C_i), i = 1, 2, \dots, n$ 是一系列按相对截止期非递减顺序排列的周期性实时消息, C_i 为 M_i 的传输时间, T_i 为消息 M_i 的周期, D_i 分别为消息 M_i 的相对截止期,且 $D_i \leq T_i$, 当满足式(4)时,对于本文接收链路调度模型,消息集 PRT 可被调度。

$$\forall (0 \leq k \leq n) \left| \left(\sum_{i=0}^k \frac{C_i}{D_i} + \frac{B_k}{D_k} \leq 1 \right) \right. \quad (4)$$

式中, $D_0 = T_0 = T_{\text{SYN}}, C_0 = C_{\text{SYN}} + \max\{C_i\}; B_0 = 0, B_i = \max\{C_i \mid D_i < D_i\}, i = 1, 2, \dots, n$ 。

证明:根据交换机对接收链路调度模型,由于时钟同步消息(SYN)可以抢占实时消息和非实时消

息的传输,而实时消息可抢占非实时消息的传输,因而时钟同步消息具有最高优先级。实时消息因可被时钟同步消息中断,最坏情况下相当于每个调度窗口均要插入长度为 $\max\{C_i\}$ 的传输空隙(图 6),为此引入虚拟周期性实时消息 $M_0 = (D_0, T_0, C_0)$ 表示时钟同步消息 SYN,其中 $D_0 = T_0 = T_{\text{SYN}}, C_0 = C_{\text{SYN}} + \max\{C_i\}$,从而得到新的消息集: $\text{PRT}_0 = \{M_i, i = 0, 1, \dots, n\}$ 。为使 $\text{PRT} = \{M_i, i = 1, 2, \dots, n\}$ 可调度,只要使 PRT_0 可调度。

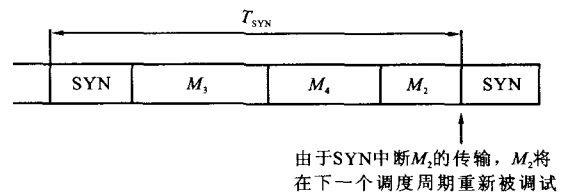


图 6 消息在接收链路上的传输

由于本文对 PRT_0 调度是基于非同步调度模型,虽然 PRT_0 中的实时消息可抢占非实时消息的传输,在实时消息集内部可能会出现高优先级的消息被低优先级所阻塞的情况,即来自不同源站点的实时消息到达同一交换机输出缓冲区的时间不同。当某一高优先级实时消息到达时,接收链路恰好被低优先级的实时消息所使用,此时只能等低优先级的实时消息完成传输后,高优先级消息才能按 EDF 算法被调度传输。显然,时钟同步消息(SYN)可以抢占实时消息和非实时消息的传输,实时消息 M_0 最坏情况下被低优先级消息阻塞时间 $B_0 = 0$;对于实时消息 $M_i, i = 1, 2, \dots, n$,在最坏情况下其被低优先级消息所阻塞的时间 $B_i = \max\{C_i \mid D_i < D_i\}$ 。根据定理 3 知,当满足式(4)时,基于非抢占式 EDF 调度算法 TS 可调度。

5 结束语

本文在文献[3]的实时消息调度模型的基础上,重点研究了实时消息在交换式以太网上的可调度判定条件,主要工作包括:(1)改进了文献[3]的实时消息可调度性分析模型,提出了一种由源站点和目标站点协同进行的分布式实时消息可调度性分析方法,简化了交换机的设计;(2)修订了文献[3]中的一些不符合实际情况的假设条件,给出了实时消息分别在发送和接收链路调度的可行性判定条件;(3)利用实时调度理论对于实时消息的可调度性判定条件进行了理论证明。

(下转第 58 页)

表 2 诊断结果比较

测试	BP 网络输出			粗糙遗传神经网络输出			期望输出		
X ₁	0.000 5	0.169 6	0.979 3	0.005 9	0.011 8	0.991 6	0	0	1
X ₂	0.000 1	0.003 4	0.814 1	0.005 3	0.002 8	0.995 3	0	0	1
X ₃	0.003 2	0.990 7	0.010 0	0.002 5	0.965 1	0.018 1	0	1	0
X ₄	0.780 1	0.998 6	0.844 0	0.990 6	0.991 4	0.996 1	1	1	1
X ₅	0.001 52	0.999 9	0.916 8	0.030 7	0.995 8	0.998 0	0	1	1
X ₆	0.000 7	0.997 7	0.999 6	0.014 9	0.999 6	0.998 5	0	1	1
X ₇	0.997 2	0.004 4	0.002 6	0.997 4	0.002 9	0.000 1	1	0	0
X ₈	0.990 0	0.010 2	0.995 5	0.997 8	0.005 4	0.998 4	1	0	1
X ₉	0.995 1	0.611 6	0.607 4	0.997 1	0.893 6	0.000 3	1	1	0

4 结束语

粗糙集、遗传算法和 BP 神经网络相结合的粗糙遗传神经网络,利用粗糙集对输入 BP 网络中的特征属性约简,去除冗余信息,应用遗传算法对 BP 网络中的权值和阈值进行优化.实践证明这种方法能够优化神经网络结构,还能对挖掘机等机械故障进行准确的识别,在工程机械故障诊断领域具有较好的应用前景.

参考文献:

[1] 刘才强,时晨,于伦正. 遗传算法在故障诊断中的应用研究[J]. 微电子学与计算机,2008,25(2):97-99.
[2] Zbigniew Michalewicz. A modified genetic algorithm for optimal control problems [J]. Computers Math Applica-

tion,1992,23(12): 83-94.

[3] Rafice J, Arvani F. Intelligent condition monitoring of a gearbox using artificial neural network[J]. Mechanical System and Signal Processing,2007,21(4): 1746-1754.
[4] 郝丽娜,王伟,吴光宇,等. 粗糙集-神经网络故障诊断方法研究[J]. 东北大学学报,2003,24(3):252-255.
[5] 王国胤. Rough 集理论与知识获取[M]. 西安:西安交通大学出版社,2001.
[6] 宋立新,王玉华,李玲远. 基于遗传算法和神经网络优化的故障诊断专家系统[J]. 华中师范大学学报:自然科学版,2005,39(3):332-335.

作者简介:

胡 平 男,(1962-),副教授.研究方向为计算智能.
彭纪奎 男,(1982-),硕士.研究方向为数据挖掘.

(上接第 54 页)

参考文献:

[1] Koeser J, Haerti H. Low-latency hard real-time communication over switched Ethernet [C]// Proc ECRTS (EUROMICRO Conf Real-Time Systems). Washington: IEEE,2004.
[2] Kweon SK, Shin KG. Achieving real-time communication over Ethernet with adaptive traffic smoothing[C]// Proceedings of the IEEE Real-Time Technology and Applications Symp. Washington: IEEE Computer Society Press, 2000: 90-100.
[3] Hoang H, Jonsson M, Hagstrom U, et al. Switched real-time ethernet with earliest deadline first scheduling protocols and traffic handling[C]// Proceedings of the International Parallel and Distributed Processing Symposium. Wisconsin: IEEE, 2002: 94-99.
[4] Devi M. An improved schedulability test for uniprocessor periodic task systems[C]// Proceedings of the 15th

Euromicro Conference on Real-Time Systems. Porto, Portugal: IEEE, 2003.

[5] 沈卓伟,汪芸. 基于 EDF 调度策略的端到端实时系统可调度性分析算法[J]. 计算机研究与发展,2006,43(5): 813-820.
[6] BAKER T P. Stack-based scheduling of real-time processes[J]. The Journal of Real-Time Systems, 1991:67-99.

作者简介:

檀 明 男,(1974-),博士,讲师.研究方向为数据库与多媒体技术、工业控制网络.
魏 臻 男,(1965-),博士,教授,博士生导师.研究方向为分布式控制、计算机控制等.
韩江洪 男,(1954-),教授,博士生导师.研究方向为分布式控制、计算机控制、计算机网络等.
钟伯成 男,(1964-),博士,教授.研究方向为计算机网络.