

区分服务网络的协作可调整 RED 算法¹

王 茜 隆克平 程时端 张润彤*

(北京邮电大学程控交换技术与通信网国家重点实验室 北京 100876)

*(Nokia 中国研发中心 和平里东街 11 号 北京 100013)

摘 要 文章描述了一种适用于区分服务网络的新的 RED 改进算法——协作可调整 RED 算法 (CARED)。算法的思想是基于令牌桶标记器的标记概率建立业务实际速率与承诺速率间的定量关系, 并以这个关系动态调整下游节点中 CARED 的丢包概率。为了实现 CARED, 文章还提出了一种标记概率的计算方法——基于包数的滑动窗口测量算法 (NSWM)。文中通过详实的仿真实验证明了定量关系的正确性, 同时验证了 CARED 算法不仅为区分服务提供了有效的拥塞控制功能, 也保障了区分服务的不同性能要求。

关键词 区分服务, 随机早期丢包 (RED), 拥塞控制, 承诺速率

中图分类号 TN919.3

1 引 言

由于 Internet 的飞速发展, 早期版本 TCP/IP 协议对拥塞控制的局限性导致了网络拥塞崩溃的现象越来越严重。多年来 IP 网络的拥塞控制一直是人们关注和研究的重点, 尤其是端到端的拥塞控制^[1]。然而, 在 Internet 网络中只有端到端的拥塞控制是不足以避免拥塞崩溃的, 网络节点也必须具有一定的拥塞缓解和控制的能力。这就是为什么随机早期丢包 (RED) 算法^[2,3]及其各种改进版本在近几年来快速发展的原因。RED 算法实际上是一种拥塞避免机制, 它通过计算排队平均队长来检测初期拥塞, 当平均队长超过预定的门限时, 随机丢弃到达的数据包。这样可以间接通知业务源减小发送速率, 直到消除了进一步的拥塞趋势。

随着 Internet 的商业化发展, 新的应用和新的服务质量需求层出不穷。而目前的 IP 网络只能提供尽力而为的服务, 它既不能支持各种各样的服务类型, 也不能保障每种服务的质量。针对这个问题, IETF 提出了区分服务体系 (简称 DiffServ)^[4,5]。尽管这种体系在一定程度上能保障不同服务类型的质量, 但拥塞仍然是降低这些服务性能的主要问题。因此, DS 网络中的路由器也需要相应的拥塞控制机制, 如 RED。但是, 各种 RED 算法的研究结果表明, RED 路由器在拥塞时对每个流的丢包概率与该流在拥塞链路上的带宽占有率近似成比例。在区分服务网络中, 如果还保持这种比例, 使不同带宽要求的服务类型还按照其占用的带宽比例来丢弃数据包, 而不考虑其特有的性能需求以及为这些需求所负担的额外费用, 这对高等级服务的用户来说是很不公平的, 特别是不能适应 AF 类服务的多类型, 多丢失优先级的特性。因此, 本文提出了一种适应区分服务网络的新的 RED 改进算法, 协作可调整 RED (CARED)。这种算法基于服务的实际速率和承诺速率间的定量关系来调整它的丢包概率。同时, 本文还对 CARED 算法的实现细节进行了讨论, 并提出了一种标记概率的计算方法——基于包数的滑动窗口测量算法 (NSWM)。最后文章通过详实的仿真验证了 CARED 算法的可行性和有效性。

2 CARED 机制

首先, 我们分析了区分服务网络边缘路由器中的令牌漏桶三色标记算法 trTCM^[6], 并利用标记概率获得各种服务的实际速率与承诺速率之间的定量关系。

¹ 2001-09-18 收到, 2001-12-31 改回

诺基亚中国研发中心合作项目和国家自然科学基金 (项目批准号 69972008) 联合资助

trTCM 通过两个令牌桶 C 和 P 来标记数据包。令牌桶 C 保存绿色令牌, 它的容量为 CBS, 令牌增长速率为 CIR; 令牌桶 P 保存黄色令牌, 它的容量为 PBS, 增长速率为 PIR。我们标识 AF 类汇聚流的到达速率为 AIRbyte/s, 平均包长为 B 字节。通过简单分析可知, 如果 $AIR \leq CIR$, 所有包都会标记成绿色, 即红包和黄包的标记概率都为零。这说明汇聚流的流量特性与其规范中的承诺速率相一致。如果 $CIR < AIR \leq PIR$, 则一部分包被标记成黄色。通过分析我们获得黄包的标记概率满足:

$$M_{yel} = \frac{E[yel]}{E[yel + gre]} = \frac{AIR - CIR}{AIR} \quad (1)$$

如果 $AIR > PIR$, 绿色和黄色令牌都处于消耗状态, 数据包可能被标记成绿色, 黄色或红色。我们同样也能得到红包的标记概率满足:

$$M_{red} = \frac{E[red]}{E[red + yel + gre]} = \frac{AIR - PIR}{AIR} \quad (2)$$

这样, 我们可以得到三色标记器中实际速率和承诺速率与标记概率的定量关系:

$$\left. \begin{aligned} \frac{AIR}{CIR} &= \frac{K}{1 - M_{red}}, & M_{red} > 0 \\ \frac{AIR}{CIR} &= \frac{1}{1 - M_{yel}}, & M_{yel} > 0 \\ \frac{AIR}{CIR} &< 1, & \text{其它} \end{aligned} \right\} \quad (3)$$

其中 $K \equiv PIR/CIR$ 且 $K \geq 1$ 。

另一方面, 有关 TCP 建模的研究结果^[1,7]表明, 如果给定网络的随机丢包概率 P , 数据包的传输周期 RTT 和平均包长 B , 一个 TCP 连接所获得的带宽上限由 (4) 式确定

$$AIR = BW \leq \frac{C \times B}{RTT \times \sqrt{P}} \quad (4)$$

其中 C 是常数。也就是说, 当网络的其它参数 (RTT 和 B) 保持相对稳定时, $AIR^2 \propto 1/P$ 。为了使 AIR 与 CIR 尽量保持一致, 可以通过调节 RED 的丢包概率 P 来调整 AIR。调节关系式为 $P \times 1/(1 - M_{yel})^2 \rightarrow P'$ 。但 $1/(1 - M_{yel})^2$ 的取值范围为 $[1, \infty)$, 调节幅度太大, 为此, 设置一个控制参数 α , 使调节因子形如 $\alpha/(1 - M_{yel})^2$, α 的取值范围是 $(0, 1]$ 。这样, 借鉴可应用于多等级服务的加权随机早期丢包 (WRED) 算法的基本思想, 基于 trTCM 标记器的 CARED 算法描述如下:

(1) trTCM 实时计算标记概率 M_{yel} 和 M_{red} 。如果 $M_{red} = 0$, 设置参数 $K = 0$; 如果 $M_{red} > 0$, 设 $K = PIR/CIR$;

(2) 标记器通过预定方式保存和传送调节因子 M 和 K , 如果 $K > 0$, 则 $M = M_{red}$ 。否则, $M = M_{yel}$;

(3) 数据包到达下游的 CARED 路由器时, 取调节因子 M 和 K ; 通过 EWMA 滤波器分别计算平均队长 avg_{gre} , avg_{yel} 和 avg_{red} ; 主动调整丢包概率如下:

$$P_{gre} = \frac{avg_{gre} - Min_{gre}}{Max_{gre} - Min_{gre}} \times Max_{p-gre}$$

当 $K = 0$ 时, 有

$$P_{yel} = \frac{avg_{yel} - Min_{yel}}{Max_{yel} - Min_{yel}} \times \frac{\alpha}{(1 - M_{yel})^2} \times Max_{p-yel}$$

其它情况时,

$$P_{yel} = \frac{avg_{yel} - Min_{yel}}{Max_{yel} - Min_{yel}} \times Max_{p-yel}; P_{red} = \frac{avg_{red} - Min_{red}}{Max_{red} - Min_{red}} \times \alpha \times \left[\frac{K}{1 - M_{red}} \right]^2 \times Max_{p-red}$$

3 CARED 的实现

CARED 与基于漏桶的标记器结合, 可以在区分服务网络中为不同类型的服务提供更有力的性能保障。但是在 DS 子网的边缘和内部路由器中实现这一机制之前, 必须先解决两个问题。一是如何实时地计量标记器的标记概率, 二是如何将调节因子从 DS 子网的边缘路由器传送到子网内的任一 CARED 路由器。

3.1 标记概率计量算法

基于令牌桶的标记概率是指某种颜色的包占所有被标记包的百分比。但是, 在一个实际网络的路由器中记录所有通过的包的个数是不现实的。因此, 我们参考了一种速率控制机制中的速率估计算法——时间滑动窗口算法 (TSW)^[8], 提出了一种标记概率计量算法——基于包数的滑动窗口计量算法 (NSWM)。具体算法如下:

- (1) 初始化: $W_{length} = \text{常数}$; $Int_{red} = Int_{yel} = Int_{gre} = 0$;
- (2) 前 W_{length} 个包标记后, $M_{red(yel, gre)} = Num_{red(yel, gre)} / W_{length}$;
- (3) W_{length} 个包之后的每个包标记后, NSWM 修改它的状态变量:

$Int_{red} ++$; $Int_{yel} ++$; $Int_{gre} ++$;

如果包标记为红色,

$$M_{red} = \frac{M_{red} \times (W_{length} + Int_{red} - 1) + 1}{W_{length} + Int_{red}}; \quad Int_{red} = 0$$

$$M_{yel} = \frac{M_{yel} \times (W_{length} + Int_{yel} - 1)}{W_{length} + Int_{yel}}; \quad M_{gre} = \frac{M_{gre} \times (W_{length} + Int_{gre} - 1)}{W_{length} + Int_{gre}}$$

如果包标记为黄色或绿色, 采用以上同样的公式修改状态, 只是标记色采用前两个公式, 非标记色采用后两个公式;

- (4) 如果在时间间隔 T_{max} 内没有包被标记, 返回 (2)

NSWM 算法记录了 8 个状态变量: $W_{length} = \text{常数}$, 表示个数滑动窗口的大小; T_{max} , 最大可不考虑的数据包的标记时间间隔; Int_{red} , Int_{yel} 和 Int_{gre} , 各种颜色包的标记数量间隔; M_{red} , M_{yel} 和 M_{gre} , 实时计量的各种颜色包的标记概率。最初我们记录 W_{length} 个标记包中每种颜色包的个数 Num_{color} , 并用这些统计数据来求出标记概率 M_{red} , M_{yel} 和 M_{gre} 的初始值, 这样可以更及时准确地反映业务量的突发特性。然后对每个新的标记包, 根据上述算法中的一系列滑动窗口公式来计量各种颜色的标记概率。

3.2 传送调节因子

标记算法和 NSWM 算法在 DS 子网的边缘路由器实现, 而 CARED 算法在边缘和内部路由器中都需要实现。因此, 我们需要将边缘标记器获得的调节因子传送到同一子网的其它路由

器中,有多种方式可以实现以上目标,如在包头的 option 域中携带这些参数,或利用 ICMP 消息传送。我们在实验中采用了如下方式:标记器中每标记一个包就调整一次状态变量,获得相应的标记概率系数。我们利用 IP 包头的 option 域来保存和传送这些标记概率系数。同时,这些参数只在每个 DS 子网内有效,当被标记包到达子网的出节点时,包头中的相关域值将被置空,避免给相邻的 DS 子网造成误解。

4 仿真及结果分析

为了验证以上的定量关系和算法,我们在 OPNET 仿真平台中进行了大量的仿真,由于篇幅有限,我们在这里只列举代表性的结果。

4.1 前提条件的验证

为了证实 NSWAM 算法的有效性和 AF 类服务实际速率与承诺速率间定量关系的正确性,我们实现了 trTCM 标记算法和 NSWAM 算法,并根据标记概率计算出汇聚流的到达速率,与仿真工具中测量到的实际速率进行了比较。

图 1 中显示的是只有黄色和绿色标记的时候, NSWAM 和定量公式计算出的速率与仿真工具测量的链路速率的比较。这两条曲线呈现非常一致的变化趋势和取值范围,可见通过标记概率计算出来的速率与实际速率是很一致的。图 2 描绘了 3 种颜色都有标记时,这两种速率曲线的比较。这两条曲线也具有同样的变化趋势,只是公式计算出的速率比测量出的速率更加平滑,抖动范围小一些。这是由于 NSWAM 算法在计算标记概率时在一定程度上平滑了概率的抖动,从而平滑了由标记概率计算出的速率。

图 1,图 2 中所示曲线的一致性充分证明了三色标记器中定量关系的正确性,也同时证实了 NSWAM 算法的有效性。以上仿真结果及分析为 CARED 算法的前提条件提供了证明。

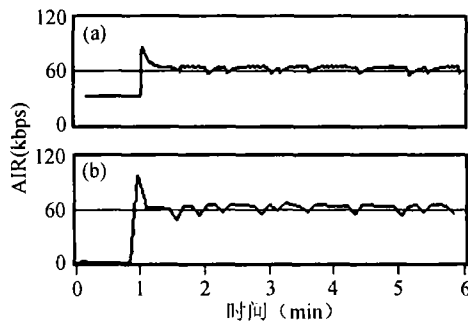


图 1 标记两色测量和计算速率的比较
(a) 计算值 (b) 测量值

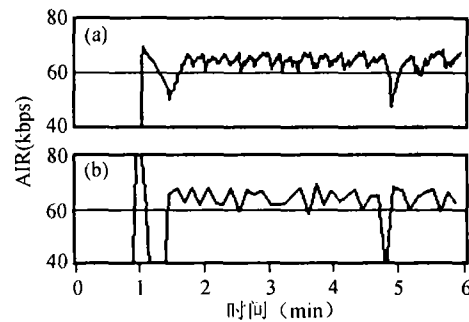


图 2 标记三色测量和计算速率的比较
(a) 计算值 (b) 测量值

4.2 CARED 的性能

为了证明 CARED 的性能优势,我们也实现了 WRED 算法。仿真中,用户分别申请不同的 AF 类服务,然后比较三类服务实际获得的速率与它们要求的承诺速率。

我们为三类服务配置了两套承诺速率,并在不同的仿真中采用。我们比较了分别采用 CARED 和 WRED 的网络中各类服务获得的速率的均值和方差。仿真网络的瓶颈速率为 1.544Mb/s,仿真中设控制参数 $\alpha = 0.5$,比较结果如表 1 所示。

表 1 两组实际获得速率与承诺速率比较

	CIR (kb/s)	CARED		WRED		CIR (kb/s)	CARED		WRED	
		均值	方差	均值	方差		均值	方差	均值	方差
AF1	400	441	411	425	444	480	487	446	537	461
AF2	200	261	282	435	471	240	303	325	392	354
AF3	200	250	276	86	218	80	160	199	23	119
总和	800	957		951		800	957		957	

从表 1 中我们可以看到相对于 WRED 来说, CARED 能够更好地保障每类 AF 服务的承诺速率, 同时还能较公平地分配剩余带宽。例如第一组比较中 AF3 类服务要求承诺速率为 200kb/s, 在 CARED 中它获得了 250kb/s 的平均带宽, 但在 WRED 中只获得了 86kb/s 的带宽, 远远小于它的承诺速率。AF2 类服务在 WRED 中获得了 435kb/s 的带宽, 甚至超过了承诺速率是其两倍的 AF1 类服务的获得带宽。因此, 仿真结果证明了 CARED 不仅保证了服务的承诺速率, 而且对剩余带宽的分配更加公平。

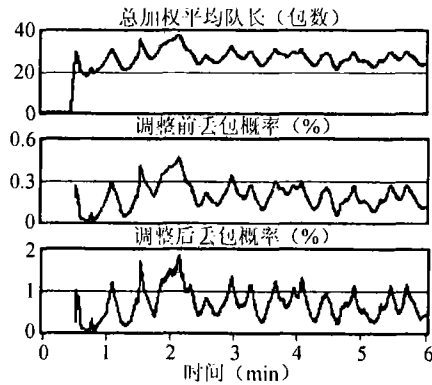


图 3 黄包丢包概率调整前后比较

另一方面, 为了证明 CARED 算法仍然具有 RED 特有的拥塞控制的能力, 我们还观察了拥塞节点中队列的总加权平均队长, 以及调整前和调整后的黄包的丢包概率。仿真结果见图 3。通过比较, 我们可以得到调节前的概率曲线实时地反映了平均队长的变化, 而调节后的概率曲线不仅能实时跟踪平均队长的变化, 还能在队长过长时进一步增大丢包概率。这也证明了, CARED 算法不仅能够实时地反映队列的拥塞程度, 控制拥塞, 还能按实际获得速率与承诺速率的一致性调节丢包概率, 充分考虑了每种服务的不同性能需求。

从以上的仿真和结果分析中, 我们可以充分证明 CARED 机制不仅可以控制区分服务网络的拥塞, 还能保障不同服务的承诺速率并较公平地分配剩余带宽, 从另一方面保障了服务的质量要求和服务间的公平性。

5 结 论

本文提出了一种适用于区分服务网络的 RED 实现机制, 即协作可调整的随机早期丢包策略 (CARED)。文中首先分析了边缘路由器中的令牌漏桶标记算法 trTCM, 并利用标记概率与实际速率和承诺速率之间的定量关系在 DS 子网路由器中动态调整不同服务类型的丢包概率。这样 CARED 不仅能够反映各类服务在总带宽中占用的份额, 还能反映这些服务占用的带宽与它们所要求的流量规范间的一致程度。大量的仿真结果证实了 CARED 算法与标记算法协作, 不仅可以依据队长占用量有效地控制拥塞, 还可以保障不同服务类型的承诺性能, 并改善各服务间的公平性。

参 考 文 献

- [1] S. Floyd, K. Fall, Promoting the use of end-to-end congestion control in the Internet, *IEEE/ACM Trans. on Networking*, 1999, 7(4), 458-472.
- [2] S. Floyd, V. Jacobson, Random early detection gateways for congestion avoidance, *IEEE/ACM Trans. on Networking*, 1993, 1(4), 397-413.
- [3] B. Braden, D. Clark, *et al.*, Recommendations on queue management and congestion avoidance in the Internet, IETF RFC2309, April 1998.
- [4] S. Blake, David Black, An architecture for differentiated services, Internet RFC 2475, October 1998
- [5] Y. Bernet, S. Blake, A framework for differentiated services, Internet draft <draft-ietf-diffserv-framework-02.txt>, Feb. 1999.
- [6] J. Heinanen, R. Guerin, A two rate three color marker, Internet RFC 2698, September 1999.
- [7] M. Mathis, J. Semke, The macroscopic behavior of the TCP congestion avoidance algorithm, *Computer Communication Review*, 1997, 27(3), 67-82.
- [8] W. Fang, N. Seddigh, B. Nandy, A time sliding window three color marker (TSWTCM). Internet RFC2859. June 2000.

COOPERATIVE ADJUSTED RED IN DiffServ NETWORK

Wang Qian Long Keping Cheng Shiduan Zhang Runtong*

(National Key Lab of Switching Tech. and Telecomm. Networks,
Beijing Univ. of Posts and Telecomm., Beijing 100876, China)

*(Nokia China R&D Center, No.11 Hepingli Dongjie, Beijing 100013, China)

Abstract This paper presents a new implementation mechanism of RED, Cooperative Adjusted RED (CARED), for the DiffServ network. It deduces first the quantitative relationship between the actual rate and the committed rate of services on the basis of the marking probability of token-based markers, and then adjust the dropping probability of CARED to drop packets fairly according to this relationship. It also proposes a new marking probability metering algorithm NSWMM to implement CARED in DiffServ network. Lots of simulations have validated the relationship between the actual rate and the committed rate and prove the efficiency of CARED.

Key words Differentiated services, Random early detection, Congestion control, Committed rate

王 茜: 女, 1975 年生, 博士生, 主要研究方向: QoS 保障的策略及实现机制、TCP/IP 协议改进等。

隆克平: 男, 1968 年生, 副教授、硕士生导师, 主要研究方向: TCP/IP 协议改进、QoS 保障的策略及其实现机制、移动 IP 技术及应用、路由器的队列调度和缓存管理策略及算法等。

程时端: 女, 1940 年生, 教授, 博士生导师, 主要研究方向: ISDN, ATM, ATM 和 IP 网的话音通信技术, 协议工程, 业务量工程, 宽带网络性能和服务质量等。