

## 基于自相似聚合业务流量的 AQM 算法性能评价\*

汪浩<sup>1,2</sup>, 严伟<sup>1+</sup>

<sup>1</sup>(北京大学 信息科学技术学院 网络实验室,北京 100871)

<sup>2</sup>(江西师范大学 软件学院,江西 南昌 330027)

### Performance Evaluation for AQM Algorithms with Self-Similar Aggregate Traffic

WANG Hao<sup>1,2</sup>, YAN Wei<sup>1+</sup>

<sup>1</sup>(Network Laboratory, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

<sup>2</sup>(School of Software, Jiangxi Normal University, Nanchang 330027, China)

+ Corresponding author: Phn: +86-10-62765811, Fax: +86-10-62765813, E-mail: yanwei@net.pku.edu.cn, http://net.pku.edu.cn

Wang H, Yan W. Performance evaluation for AQM algorithms with self-similar aggregate traffic. *Journal of Software*, 2006,17(9):1958-1968. <http://www.jos.org.cn/1000-9825/17/1958.htm>

**Abstract:** The existing TCP/AQM model does not consider the impact of unresponsive flows to AQM algorithms, which contributes to about 70%~80% of the Internet flows. It is important to analyze the performance of AQM algorithms taking into account of the unresponsive flows. An extended GI/M/1/N queueing system is established by means of embedding the AQM mechanism into the standard GI/M/1/N queueing system. Based on the extended GI/M/1/N queueing system and self-similar traffic of the Internet, three classical AQM algorithms (TD, RED and GRED) are evaluated for the unresponsive flows. The analytic results are consistent with those obtained from NS2 simulations, which means the extended queueing system can be used to assess the performance of AQM algorithms with the unresponsive flows.

**Key words:** computer network; active queue management; congestion control; GI/M/1/N queueing system; heavy tailed distribution

**摘要:** 现有 TCP/AQM 忽略了非响应业务流量对 AQM 算法性能的影响,但非响应业务流量约占 Internet 业务流量的 70%~80%。因此,评价非响应业务流量对 AQM 算法性能的影响具有重要意义。借助于在标准 GI/M/1/N 排队系统中嵌入 AQM 算法随机丢包机制的手段,提出了一种利用“扩充的 GI/M/1/N 排队系统”评价 AQM 算法在非响应业务流量下的性能的分析方法。最后评价了 TD、RED 和 GRED 这 3 种经典的 AQM 算法,评价结果与 NS-2 模拟结果一致,表明该分析方法可能用于评价 AQM 算法在非响应业务流量下的性能。

**关键词:** 计算机网络;主动队列管理(AQM);拥塞控制;GI/M/1/N 排队系统;重尾分布

中图分类号: TP393 文献标识码: A

由于尾丢包 TD(tail drop)不能在网络出现拥塞的征兆时及时通知 TCP 主机降低发送 IP 包的速率,因此,

\* Supported by the National Grand Fundamental Research 973 Program of China under Grant No.2004CB318204 (国家重点基础研究发展规划(973)); the Cisco Academic Research Project of China (思科教育科研资助项目)

Received 2005-03-21; Accepted 2006-01-24

IETF 提出了主动队列管理(AQM),其作用是当路由器发现网络出现拥塞的征兆时,按一定的概率随机丢弃到达的 IP 包,以便通知 TCP 主机降低发送 IP 包的速率.IETF 在 RFC 2309 中推荐 RED<sup>[1]</sup>为唯一的 AQM 候选算法.Internet 业务流量由响应业务流量(即长效 TCP 业务流量)和非响应业务流量(即短效 TCP 业务流量,如 Web 业务流量和 UDP 业务流量)构成.当 AQM 算法的丢包信号到达短效 TCP 业务流量的源端时,短效 TCP 业务流量的源端通常已无数据可发送,故短效 TCP 业务流量不能响应 AQM 算法的丢包信号.UDP 流本身不具备拥塞控制机制,故也不能响应 AQM 的丢包信号.研究表明,Internet 流量主要由短效 TCP 业务流量(如 Web 流量)构成,其中 65%~80%的字节流量属于短效 TCP 流量;55%~75%的 IP 包属于短效 TCP 流量<sup>[2]</sup>.此外,2004 年度 Sigcomm 获奖者 Simon 认为:随着分布式多媒体应用的普及,越来越多的 UDP 流将会出现在 Internet 上,并对 Internet 的稳定性产生重要影响<sup>\*</sup>.因此,评价非响应业务流量对 AQM 算法性能的影响不仅具有重要意义,而且具有前瞻性.

现有 TCP/AQM 模型<sup>[3,4]</sup>重点放在分析 AQM 与响应业务流量相互作用的反馈机制上,为用控制论方法设计新的 AQM 算法提供理论基础<sup>[5]</sup>.但是,TCP/AQM 模型忽略了非响应业务流量对 AQM 算法性能的影响.因为非响应业务流量的发送端不能响应 AQM 算法的丢包信号,所以可利用开环的排队系统评价 AQM 算法的性能.为此,我们在标准的 GI/M/1/N 排队系统中嵌入了 AQM 算法的随机丢包机制,建立了一个“扩充的 GI/M/1/N 排队系统”,并提出了一种利用“扩充的 GI/M/1/N 排队系统”评价 AQM 算法在非响应业务流量下的性能的分析方法.利用该方法,我们评价了 TD,RED 和 GRED 这 3 种经典的 AQM 算法,评价结果与其他研究者利用模拟或实验方法所得到的结果一致<sup>[6-8]</sup>.为了进一步验证该分析方法的合理性,我们利用 NS-2 进行了一系列的仿真实验.仿真实验结果表明,该分析方法所得结果与仿真实验所得结果一致.因此,该分析方法可用于评价 AQM 算法在非响应业务流量下的性能.

该方法具有如下优越性:1) 能够更真实地刻画 AQM 算法在 Internet 业务流量下的性能;2) 无须设计模拟或实验场景,避免了主观因素对 AQM 算法性能评价的干扰.

#### 相关工作

到目前为止,大多数研究人员借助网络模拟或网络实验方法来比较和评价不同 AQM 算法的性能<sup>[6-8]</sup>:Brandauer 利用模拟和实验手段比较了 TD,RED 和 GRED 算法在非响应业务流量下的性能<sup>[7]</sup>;Iannaccone 利用实验手段比较了 TD,RED 和 GRED 算法在聚合业务流量下的性能<sup>[8]</sup>;Eitan 等人通过模拟手段分析了短效 TCP 业务流量条件下 RED 算法的公平性<sup>[9]</sup>;Christiansen 等人通过实验手段分析了 RED 算法对 Web 请求响应时间的影响<sup>[10]</sup>.

由于非响应业务流量不能响应 AQM 算法的丢包信号,故另有一些研究者试图用开环的排队论模型评价 AQM 算法的性能:Bonald 利用 M/M/1/N 排队系统比较了 TD 和 RED 算法在 Poisson 业务流量下的性能<sup>[11]</sup>;Garetto 利用 M<sup>X</sup>/M/1/N 排队系统分析了路由器队列长度(时延)的分布<sup>[12]</sup>;Younsuk 利用 GI/M/1/N 排队系统在 IP 包的到达间隔服从 Pareto 分布时,讨论了 TD 算法的丢包率<sup>[13]</sup>;Alazemi 利用二维二阶 Markov 链分析了 RED 算法在 Poisson 业务流量下的性能<sup>[14]</sup>.

为了用排队模型评价非响应业务流量条件下 AQM 算法的性能,需知道 IP 包到达路由器间隔时间的分布规律.网络测量的统计结果表明,IP 包的到达间隔时间服从重尾 Weibull 分布或 Pareto 分布,并且趋于相互独立和指数分布<sup>[15-17]</sup>.因此,可以利用这些分布作为排队模型的输入,以评价非响应业务流量对 AQM 算法性能的影响.

#### 本文的贡献

本文提出了一种利用“扩充的 GI/M/1/N 排队系统”评价 AQM 算法在非响应业务流量下的性能的分析方法.同时,分析评价了 TD,RED 和 GRED 这 3 种经典的 AQM 算法,评价结果不仅与我们利用 NS-2 的仿真实验结果一致,同时也与其他研究者利用模拟或实验方法所得结果一致<sup>[7,8]</sup>,从而表明该分析方法可用于评价 AQM 算法的性能.

本文第 1 节讨论扩充的 GI/M/1/N 排队系统.第 2 节讨论评价 AQM 算法性能的指标体系.第 3 节评价比较

\* Keynote Speech by SIGCOMM Award Winner Simon Lam. [http://www.acm.org/sigs/sigcomm/sigcomm2004/conf\\_program.html](http://www.acm.org/sigs/sigcomm/sigcomm2004/conf_program.html), 2004.

TD,RED 和 GRED 算法的性能.第 4 节利用 NS-2 仿真实验验证利用扩充的 GI/M/1/N 排队系统评价 AQM 算法性能的合理性.第 5 节是总结.

### 1 扩充的 GI/M/1/N 排队系统

从本节开始,我们约定如下记号: $N$  表示路由器能容纳的 IP 包的个数,其中的一个 IP 包正在被转发,其余的 IP 包( $\leq N-1$ )在缓存队列等待处理; $df$  表示概率分布函数; $pdf$  表示概率密度函数; $r.v.$  表示一个随机变量; $r.v.s$  表示多个随机变量; $i.i.d$  表示独立同分布.

#### 1.1 扩充GI/M/1/N排队系统的基本思路

如果缓存队列未滿,标准的 GI/M/1/N 排队系统将接纳到达的每个 IP 包,并放在缓存队列中(如图 1 所示).此时,我们可以直接利用文献[18]的结果计算队列长度的概率分布,并据此计算丢包率、平均队列长度、队列长度的方差等性能指标.标准的 GI/M/1/N 排队系统仅能用于评价 TD 算法的性能.

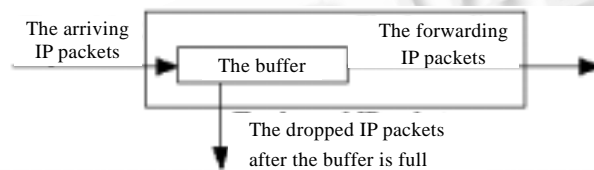


Fig.1 A standard GI/M/1/N queueing system

图 1 标准的 GI/M/1/N 排队系统

在嵌入 AQM 算法的丢包机制后,即使缓存队列未滿,排队系统仍然会随机地丢弃 IP 包,从而得到扩充的 GI/M/1/N 排队系统(如图 2 所示).

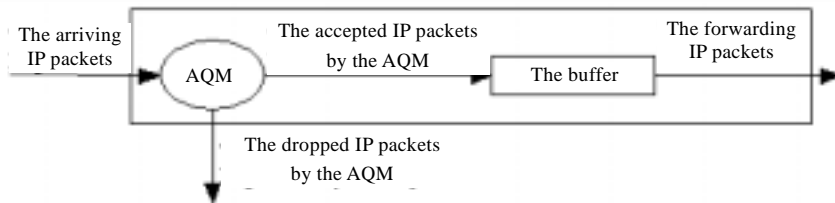


Fig.2 An extended GI/M/1/N queueing system

图 2 扩充的 GI/M/1/N 排队系统

假设到达排队系统的 IP 包到达间隔的  $df$  是  $A(t)$ ,由于 AQM 算法随机丢弃到达排队系统的 IP 包,故真正进入缓存队列的 IP 包变得“稀疏”了.假设到达缓存队列的 IP 包到达间隔的  $df$  为  $A'(t)$ ,一般而言, $A'(t) \neq A(t)$ ,且  $A'(t)$  依赖于 AQM 算法的丢包概率和  $A(t)$ .

因此,扩充 GI/M/1/N 排队系统的基本思路如下:

1) 根据 AQM 算法的丢包概率和  $A(t)$ ,求解  $A'(t)$ ,得到定理 1;

2) 利用排队论中的辅助变量法(参见文献[19]第 296~308 页)和 Markov 嵌入链,求解队列长度的概率分布,得到定理 2 和定理 3.

我们称满足条件  $0=d_0 \leq d_1 \leq \dots \leq d_{N-1} < d_N=1$  的序列  $\{d_i\} (i=0,1,\dots,N)$  为丢包概率序列,而称序列  $\{r_i; r_i=1-d_i, i=0,1,\dots,N\}$  为收包概率序列.换言之,假设路由器中当前有  $i$  个 IP 包,当一个新的 IP 包到达时,AQM 算法以概率  $d_i$  丢弃该 IP 包,或者说以概率  $r_i$  接收该 IP 包.例如,RED 的丢包概率序列为(这里我们采用文献[11]的方法,用瞬时队列长度计算丢包概率,并且令  $\max_{th}=N$ )

$$\{d_i; d_i=0, \text{ if } 0 \leq i < \min_{th}; d_i = \frac{i - \min_{th}}{\max_{th} - \min_{th}} \max_p, \text{ if } \min_{th} \leq i < \max_{th}; d_i=1, \text{ if } i \geq \max_{th}\}.$$

## 1.2 扩充的GI/M/1/N排队系统

假设 IP 包到达路由器的间隔时间是 i.i.d r.v.s,其 df 是  $A(t)$ ,pdf 是  $a(t)$ ,均值为  $a, t \geq 0$ .如果路由器中当前有  $i$  ( $i=0,1,\dots,N$ )个 IP 包,当一个 IP 包到达时,该 IP 包以概率  $r_i$  进入路由器缓存队列,或者说以概率  $d_i$  被路由器(AQM 算法)丢弃,其中  $r_i+d_i=1$ .路由器按先到先服务的方式转发 IP 包,转发时间服从均值为  $1/\mu$ 的指数分布,且与到达间隔时间相互独立.此排队系统的输入强度为  $\rho = \frac{1}{a\mu}$ .

求解排队系统的关键步骤是计算队列长度的概率分布.尽管我们已知 IP 包到达路由器的间隔时间的 df 是  $A(t)$ ,但由于丢包的原因,IP 包到达路由器缓存队列的间隔时间的 df 已不再是  $A(t)$ .我们约定路由器中当前有  $i$  ( $i=0,1,\dots,N$ )个 IP 包,且当丢包概率为  $d_i$  时,IP 包到达路由器缓存队列的间隔时间的 df 为  $A_i(t)$ ,pdf 为  $a_i(t)$ ,均值为  $a_i, t \geq 0$ ,如图 3 所示.我们的目标是要找到用  $d_i$  和  $A(t)$  表示  $A_i(t)$  或  $a_i(t)$  的方法.

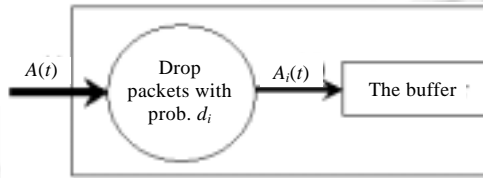


Fig.3 The thinning of input flows

图 3 输入流的稀疏化

如果以概率  $r_i$  按几何分布稀疏输入流(对应 AQM 以概率  $d_i$  按几何分布丢包),则  $A_i(t)$  可以通过  $r_i$  (或  $d_i$ ) 和  $A(t)$  表示为(参见文献[19]第 68 页)

$$A_i(t) = \sum_{l=1}^{\infty} r_i (1-r_i)^{l-1} A_i^l(t) = \sum_{l=1}^{\infty} (1-d_i) d_i^{l-1} A_i^l(t) \quad (1)$$

其中,  $A_i^l(t)$  表示  $A(t)$  的  $l$  次卷积运算.在 Weibull, Pareto 和指数分布中,只有指数分布函数的卷积能够给出简单的解析式;对 Weibull, Pareto 等重尾分布函数(更准确地,应称为 subexponential 分布函数,它是重尾分布函数集的一个子集),可以根据重尾分布函数的性质得到公式(1)的近似表达式,参见文献[20].为此,我们证明了如下定理:

**定理 1.** 如果  $A(t)$  是属于 subexponential 的分布函数<sup>[20]</sup>,离散型 r.v.  $v_i$  为路由器缓存队列接受一个 IP 包的间隔次数(如附录 1 中图 17 所示),则

i) 对  $i=0,1,\dots,N-1, A_i(t)$  可用下式估计:

$$A_i(t) = 1 - E[v_i][1 - A(t)] \quad (2)$$

其中,  $t \geq \tau$  且  $\tau$  满足  $\int_{\tau}^{\infty} a(t) dt = \frac{1}{E[v_i]}$ . 如果以概率  $r_i$  按几何分布稀疏输入流(对应 AQM 以概率  $d_i$  按几何分布丢包),

则  $E[v_i] = \frac{1}{r_i}$ ; 如果以概率  $r_i$  按均匀分布稀疏输入流(对应 AQM 以概率  $d_i$  按均匀分布丢包),则  $E[v_i] \approx \frac{1}{2} + \frac{1}{2r_i}$ ;

ii) 对  $i=N, A_N(t)$  可用下式估计:

$$A_N(t) = 1 - e^{-\frac{t}{a}} \quad (3)$$

定理 1 的证明将在附录 1 中给出.我们下一步将讨论如何求解“扩充的 GI/M/1/N 排队系统”的队列长度的概率分布.先引入两个记号:  $q_k$  ( $k=0,1,\dots,N$ ) 表示当一个新包到达时,路由器缓存队列中有  $k$  个包的概率;  $p_k$  ( $k=0,1,\dots,N$ ) 表示任意时刻路由器缓存队列中有  $k$  个包的概率.

我们首先用排队论中的辅助变量法(参见文献[18]或文献[19]的第 296~308 页)找出  $q_k$  和  $p_k$  之间的关系;然后用 Markov 嵌入链的方法求解  $q_k$ .为此,我们证明了如下定理<sup>[21]</sup>:

**定理 2.** 如果已知  $q_k$  ( $k=0,1,\dots,N$ ),则计算  $p_k$  的公式为

$$p_{k+1} = \frac{1}{\mu a_k} q_k, \quad k=0,1,\dots,N-1 \quad (4)$$

$$p_0 = 1 - \sum_{k=1}^N p_k \quad (5)$$

其中,  $a_k$  为以概率  $d_k$  丢包时, 分布函数  $A_k(t)$  的均值.

定理 3.  $q_k$  可以通过求解以下线性方程组获得:

$$q_j = \sum_{i=0}^N q_i h_{i,j}, \quad j=0, 1, \dots, N \quad (6)$$

$$\sum_{j=0}^N q_j = 1 \quad (7)$$

其中, 计算该线性方程组系数  $h_{i,j}$  的公式为

$$h_{i,j} = \int_{\tau}^{\infty} \frac{e^{-\mu t} (\mu t)^{i+1-j}}{(i+1-j)!} dA_i(t), \quad 0 \leq i \leq N-1, \quad 1 \leq j \leq i+1 \quad (8)$$

$$h_{i,0} = 1 - \sum_{j=1}^{i+1} h_{i,j}, \quad 0 \leq i \leq N-1 \quad (9)$$

$$h_{N,j} = \int_0^{\infty} \frac{e^{-\mu t} (\mu t)^{N-j}}{(N-j)!} dA_N(t), \quad 1 \leq j \leq N \quad (10)$$

$$h_{N,0} = 1 - \sum_{j=1}^N h_{N,j} \quad (11)$$

一旦给定 IP 包到达路由器的间隔时间的分布函数和 AQM 算法, 根据定理 1~定理 3, 我们就能获得路由器队列长度的概率分布, 并据此计算第 2 节中的评价 AQM 算法性能的指标.

## 2 评价 AQM 算法性能的指标体系

借助第 1 节得到的队列长度的概率分布  $q_k$  和  $p_k$ , 本节将给出丢包率、链路利用率、队列长度的均值和方差等 AQM 算法性能评价指标的计算公式.

如果用 LossRate 表示 AQM 算法的丢包率, 根据全概率公式, 得到

$$\text{LossRate} = \sum_{k=0}^N d_k q_k \quad (12)$$

链路利用率表示链路不空闲的概率. 注意到,  $p_0$  表示任意时刻路由器缓存队列为空 (有 0 个 IP 包) 的概率; 因此, 如果用 LinkUtili 表示 AQM 算法的链路利用率, 则

$$\text{LinkUtili} = 1 - p_0 \quad (13)$$

通常用队列长度的均值度量延迟, 用方差度量延迟抖动. 如果用随机变量  $\xi$  表示任意时刻路由器缓存队列的长度 (需要减去正在被转发的 IP 包), 则队列长度的均值和方差的计算公式为

$$\bar{\xi} = E[\xi] = \sum_{k=1}^N (k-1) p_k \quad (14)$$

$$\text{Var}[\xi] = \sum_{k=1}^N (k-1-\bar{\xi})^2 p_k \quad (15)$$

## 3 TD, RED 和 GRED 算法的性能评价

根据扩充的 GI/M/1/N 模型和 AQM 算法性能评价指标, 借助数值分析的方法, 我们能获得 AQM 算法的性能评价指标, 具体步骤如下:

- 1) 对给定的 AQM 算法, 计算丢包概率序列  $\{d_i\} (i=0, 1, \dots, N)$ ;
- 2) 对给定的 IP 包到达间隔时间分布函数 (Weibull, Pareto), 根据定理 1 确定  $A_i(t)$ ; 然后根据式 (8)~式 (11) 计算  $h_{i,j}$ ;

3) 根据式 (6)、式 (7) 求解  $q_k (k=0, 1, \dots, N)$ ;

4) 根据定理 2 计算  $p_k (k=0, 1, \dots, N)$ ;

5) 根据式 (12)~式 (15) 计算性能评价指标, 并据此评价 AQM 算法的性能.

本节利用扩充的 GI/M/1/N 排队系统 (见第 1 节) 及 AQM 算法性能评价指标 (见第 2 节) 评价了 3 种经典 AQM 算法 TD, RED 和 GRED 的性能. 按表 1 的参数设置, 我们对路由器缓存队列为  $N=50$  的情况, 计算了 TD, RED 和 GRED 算法在 IP 包的到达间隔分别服从 Pareto, Weibull 和指数分布, 以及不同的输入强度下的性能指标. 我们先

给出 TD,RED 和 GRED 算法在到达间隔时间为 Pareto 分布时的性能曲线,然后给出 RED 算法在到达间隔时间分别为 Pareto,Weibull 和指数分布时的性能曲线.图例中的 TD\_S,RED\_S 和 GRED\_S 分别表示 TD,RED 和 GRED 算法.

Table 1 The settings of parameters of the compared AQMs

表 1 算法的参数设置

	N=50		
	max <sub>p</sub>	min <sub>th</sub>	max <sub>th</sub>
TD	NA	NA	NA
RED	0.1	15	50
GRED	0.1	15	35

3.1 Pareto到达间隔下,TD,RED和GRED的性能比较

通过观察图 4 中 TD,RED 和 GRED 算法的丢包率的折线图,可以发现:

- 1) 随着输入强度的增加,TD,RED 和 GRED 算法的丢包率逐渐增加;
- 2) 在不同的输入强度下,尽管 TD 的丢包率<RED 的丢包率<GRED 的丢包率,但 TD 的丢包率、RED 的丢包率和 GRED 的丢包率之间没有明显的区别,也就是说,与 TD 相比,GRED 和 RED 不能明显地降低丢包率.这一结论也被许多模拟和实验结果所证实<sup>[6,8]</sup>.

通过观察图 5 中 TD,RED 和 GRED 算法的链路利用率的折线图,可以发现:

- 1) 随着输入强度的增加,TD,RED 和 GRED 算法的链路利用率逐渐增加;
- 2) 随着输入强度的增加,尽管 TD 的链路利用率>RED 的链路利用率>GRED 的链路利用率,但 TD 的链路利用率、RED 的链路利用率和 GRED 的链路利用率之间没有明显区别.这一结论也被许多模拟和实验结果所证实<sup>[6,8]</sup>.

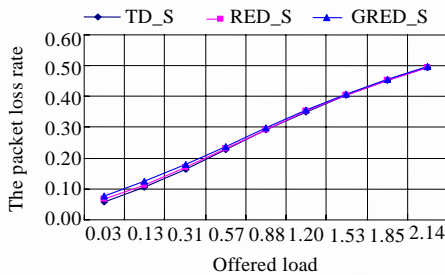


Fig.4 The packet lose rate

图 4 丢包率

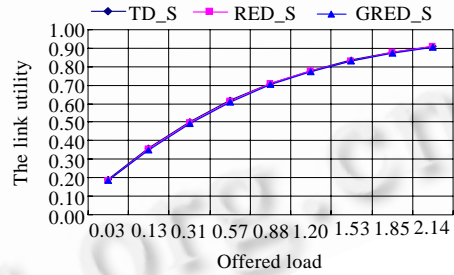


Fig.5 The link utility

图 5 链路利用率

平均队列长度是延迟的一种度量.通过观察图 6 中 TD,RED 和 GRED 算法平均队列长度的折线图可以发现:

- 1) 随着输入强度的增加,TD,RED 和 GRED 算法的平均队列长度逐渐增加;
- 2) 在不同的输入强度下,GRED 的平均队列长度<RED 的平均队列长度<TD 的平均队列长度.

该结果与文献[6,8]的结果一致.

类似地,队列长度的方差是延迟抖动的一种度量.通过观察图 7 中 TD,RED 和 GRED 算法的队列长度方差的折线图,可以发现:

1) 随着输入强度的增加,当输入强度小于 1 时,TD,RED 和 GRED 算法队列长度的方差逐渐增加;当输入强度大于 1 时,TD,RED 和 GRED 算法队列长度的方差逐渐减小.这表明:当输入强度逐渐增加但小于 1 时,由于平均队列长度较小,瞬时队列长度围绕平均队列长度振荡的幅度逐渐增大;当输入强度逐渐增加且大于 1 时,由于平均队列长度较大(接近 N=50),瞬时队列长度围绕平均队列长度振荡的幅度逐渐减小;

- 2) TD,RED 和 GRED 算法的队列长度的方差有一些差异.当输入强度小于 1 时,TD 的方差>RED 的方差>

GRED 的方差,但差异不大;当输入强度大于 1 时,TD 的方差<RED 的方差<GRED 的方差,差异也不大.

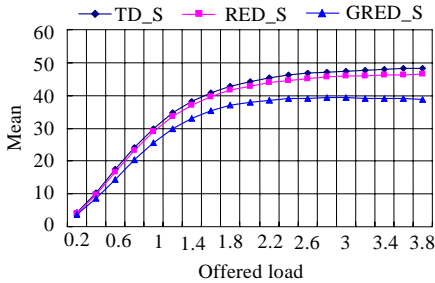


Fig.6 The mean of queue length  
图 6 平均队列长度

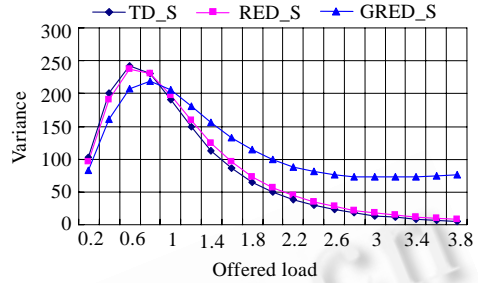


Fig.7 The variance of queue length  
图 7 队列长度的方差

### 3.2 Pareto,Weibull和指数到达间隔下的RED性能比较

图 8 和图 9 分别表示 RED 算法在 Pareto,Weibull 和 Exponent(指数)到达间隔下的丢包率和链路利用率.从图中可知,Pareto 具有最大丢包率和最小链路利用率;Exponent 具有最小丢包率和最大链路利用率;Weibull 夹在它们中间.Weibull 和 Exponent 仅当输入强度在 0.4~1.2 之间时才相互分离;而当输入强度小于 0.4 或大于 1.2 时,Weibull 和 Exponent 相互重合.注意到:当  $1 < \alpha \leq 2$  时,Pareto 分布具有无穷方差.这意味着,当网络业务流量的突发性增大时,丢包率也随之增大.

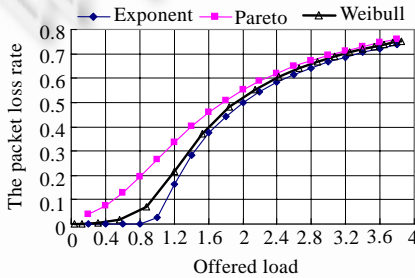


Fig.8 The packet loss rate of RED  
图 8 RED 算法的丢包率

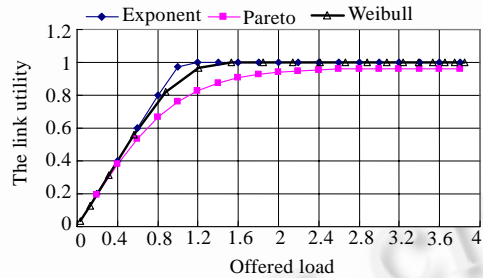


Fig.9 The link utility of RED  
图 9 RED 算法的链路利用率

图 10 和图 11 分别表示 RED 算法在 Pareto,Weibull 和 Exponent (指数)到达间隔下队列长度的均值和方差.从图 10 可知,当输入强度小于 1 时,Pareto 具有最大的平均队列长度;Exponent 具有最小的平均队列长度;Weibull 夹在它们中间.当输入强度大于 1 时,Pareto 具有最小的平均队列长度;Exponent 具有最大的平均队列长度;Weibull 夹在它们中间.

从图 11 可知,Pareto 队列长度的方差最大;Exponent 最小;Weibull 夹在它们中间.当输入强度在 0.8~1.2 之间时,队列长度的方差达到峰值,此时路由器缓存队列振荡幅度最大.

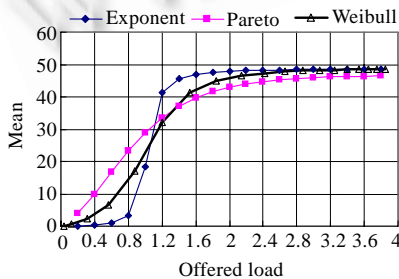


Fig.10 The mean of queue length of RED  
图 10 RED 算法的平均队列长度

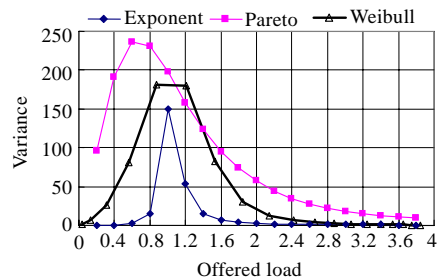


Fig.11 The variance of queue length of RED  
图 11 RED 算法队列长度的方差

## 4 仿真实验结果

为了验证扩充的 GI/M/1/N 排队系统用于评价 AQM 算法在非响应业务流量下性能的合理性,我们用 NS-2 做了一系列的仿真实验,并与第 3 节用分析方法所得结果进行了对比。

仿真实验采用与文献[9]相同的哑铃型网络拓扑(如图 12 所示)及网络参数.注意到:在做仿真实验时,瓶颈链路 N-D 上“IP 包丢失率和链路利用率”对应图 2 中扩充的 GI/M/1/N 排队系统的“IP 包丢失率和链路利用率”。

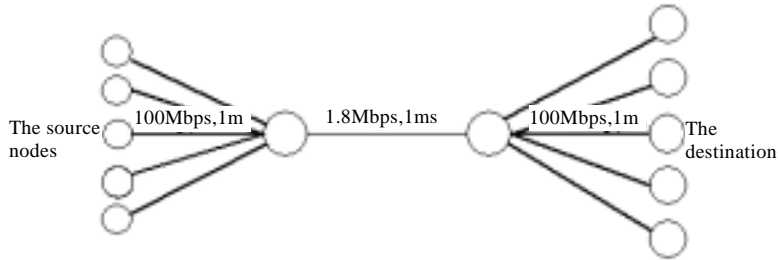


Fig.12 The topology of network

图 12 网络拓扑结构

仿真实验采用的哑铃型网络拓扑描述如下:若干个源节点通过瓶颈链路 N-D 与若干个目标节点相连,源节点中有若干个发送 TCP 业务流量的 TCP 节点和发送 UDP 业务流量的 UDP 节点.瓶颈链路 N-D 的带宽为 1.8Mbps,传播时延为 1ms,队列长度为 50 packets;其他链路的带宽均为 100Mbps,传播时延为 1m.为了仿真实际的 Web 流量(采用与文献[9]相同的方法),每个 TCP 节点依次发起 100 个短效 TCP 会话,短效 TCP 会话之间的发起间隔时间服从负指数分布,平均间隔时间为 0.9 秒;并且每个 TCP 会话传输的文件大小服从 Pareto 分布,平均大小为 10kb~20kb(这也是 Web 页面的平均大小).TCP 客户端使用 Reno 或 NewReno.UDP 节点从模拟开始至结束,一直发送 CBR 流量.通过控制 TCP 节点和 UDP 节点的数量,我们可以生成具有不同输入强度的业务流量.其中,TCP 节点产生的短效业务流量占 80%~90%;UDP 节点产生的业务流量占 10%~20%.为了便于与第 3 节分析方法所得结果相比较,仿真实验中 TD,RED 和 GRED 的参数设置与第 3 节相同.表 2 是各次仿真实验中其他一些参数的设置.

Table 2 The settings of parameters of the simulations

表 2 各次仿真实验的参数设置

TCP's node number.	UDP's load number	The averaging load (kb)	TCP's client
3	1	10	Reno
20	4		
50	6	15	
80	9		
100	12	20	Newreno
120	13		

针对不同输入强度的业务流量(共 6 种不同强度的业务流量)、被传输文件的平均大小(共 3 种不同的平均大小)、TCP 客户端(共 2 种不同的客户端)和不同的 AQM 算法(TD,RED 和 GRED),我们总共做了  $6 \times 3 \times 2 \times 3 = 108$  次模拟实验.通过分析仿真结果,我们获得了 IP 包丢失率、链路利用率、队列长度的均值与方差等 AQM 算法的性能指标.下面是部分仿真实验的结果.

图 13 和图 14 分别为被传输文件平均大小=10kb,TCP 客户端为 Newreno 时,仿真实验所得到的相应算法的 IP 包丢包率和链路利用率.图 15 和图 16 分别为被传输文件平均大小为 20kb,TCP 客户端为 Reno 时,仿真实验所得到的相应算法的 IP 包丢包率和链路利用率.



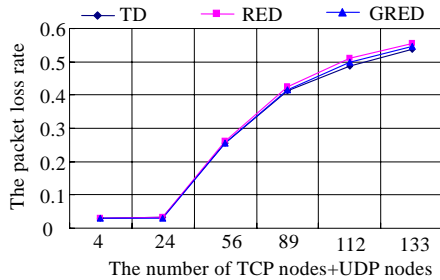


Fig.13 The packet lose rate (10kb, Newreno)

图 13 丢包率(10kb,Newreno)

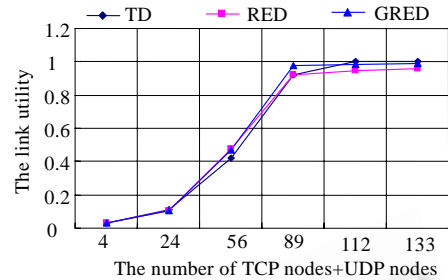


Fig.14 The link utility (10kb, Newreno)

图 14 链路利用率(10kb,Newreno)

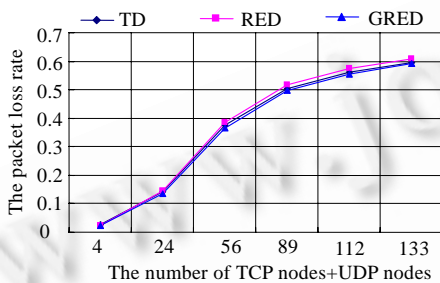


Fig.15 The packet lose rate (20kb, Reno)

图 15 丢包率(20kb,Reno)

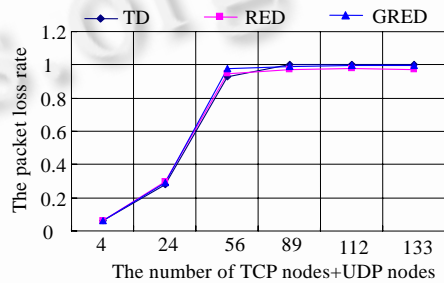


Fig.16 The link utility (20kb, Reno)

图 16 链路利用率(20kb,Reno)

通过观察图 13 和图 15 中 TD,RED 和 GRED 算法的 IP 包丢包率的折线图可以发现:随着输入强度(TCP 节点数+UDP 节点数)的增加,TD,RED 和 GRED 算法的 IP 包丢包率逐渐增加;但 TD,RED 和 GRED 算法的丢包率之间没有明显区别.该仿真实验的结果与图 4 利用分析方法所得结论一致,也与文献[6,8]的结论一致.

通过观察图 14 和图 16 中 TD,RED 和 GRED 算法的链路利用率的折线图可以发现:随着输入强度的增加,TD,RED 和 GRED 算法的链路利用率逐渐增加,但 TD,RED 和 GRED 算法的链路利用率之间没有明显区别.该仿真实验的结果与图 5 利用分析方法所得结果一致,也与文献[6,8]的结论一致.

## 5 总结

由于 TCP/AQM 模型忽略了 Internet 非响应业务流量对 AQM 算法性能的影响,而非响应业务流量占 Internet 业务流量的 70%~80%,因此,本文提出了一种利用扩充的 GI/M/1/N 排队系统来评价 AQM 算法在非响应业务流量下的性能的分析方法,并且评价比较了 TD,RED 和 GRED 等 3 个经典的 AQM 算法.为了验证该方法的有效性,我们用 NS-2 进行了一系列的模拟实验,实验所得结果与模型的分析结果相一致.此外,模型的分析结果与其他研究者的模拟和实验结果也一致.以上事实表明,扩充的 GI/M/1/N 排队系统能用于评价 AQM 算法在非响应业务流量下的性能.

致谢 衷心感谢审稿专家给予本文的有益建议和帮助.

## References:

- [1] Floyd S, Jacobson V. Random early detection gateways for congestion avoidance. IEEE/ACM Trans. on Networking, 1993,1(4): 397-413.
- [2] Thompson K, Miller GJ, Wilder R. Wide-Area Internet traffic patterns and characteristics. IEEE Network, 1997,11(6):10-23.

- [3] Misra V, Gong W, Towsley D. Fluid-Based analysis of a network of AQM routers supporting TCP flows with an application to RED. *ACM SIGCOMM Computer Communication Review*, 2000,30(4):151–160.
- [4] Tinnakornsrisuphap P, Richard JL. Characterization of queue fluctuations in probabilistic AQM mechanisms. *ACM SIGMETRICS Performance Evaluation Review*, 2004,32(1):283–294.
- [5] Hollot CV, Misra V, Towsley D, Gong WB. On designing improved controllers for AQM routers supporting TCP flows. In: *Proc. of the IEEE INFOCOM 2001*. IEEE, 2001. 1726–1734.
- [6] Trinh AT, Molnar SA. A comprehensive performance analysis of random early detection mechanism. *Telecommunication Systems*, 2004,25(1-2):9–31.
- [7] Brandauer C, Iannaccone G, Diot C, Ziegler T, Fdida S, May M. Comparison of tail drop and active queue management performance for bulk-data and Web-like Internet traffic. In: *Proc. of the 6th IEEE Symp. on Computers and Communications*. IEEE, 2001. 122–129.
- [8] Iannaccone G, May M, Diot C. Aggregate traffic performance with active queue management and drop from tail. *ACM SIGCOMM Computer Communication Review*, 2001,31(3):4–13.
- [9] Eitan A, Tania J. Simulation analysis of RED with short lived TCP connections. *Computer Network*, 2004,44(5):631–641.
- [10] Christiansen M, Jeffay K, Ott D, Smith FD. Tuning RED for web traffic. *IEEE/ACM Tran. on Networking*, 2001,9(3):249–264.
- [11] Bonald T, May M, Bolot JC. Analytic evaluation of RED performance. In: *Proc. of the IEEE INFOCOM 2000*. IEEE, 2000. 1415–1424.
- [12] Garetto M, Towsley D. Modeling, simulation and measurements of queuing delay under long-tail Internet traffic. *ACM SIGMETRICS Performance Evaluation Review*, 2003,31(1):47–57.
- [13] Younsuk K, Kiseon K. Loss probability behavior of Pareto/M/1/K queue. *IEEE Communication Letters*, 2003,7(1):39–41.
- [14] Alazemi HM, Mokhtar A, Azizoglu M. Stochastic approach for modeling random early detection gateways in TCP/IP networks. In: *Proc. of the IEEE ICC 2001*. IEEE, 2001. 2385–2390.
- [15] Cao J, William SC, Lin D, Don XS. Internet traffic tends toward poisson and independent as the load increase. In: Holmes C, ed. *Proc. of the Nonlinear Estimation and Classification*. New York: Springer-Verlag, 2002. 1–18.
- [16] Cao J, William SC, Lin D, Don XS. On the nonstationarity of Internet traffic. *ACM SIGMETRICS Performance Evaluation Review*, 2001,29(1):102–112.
- [17] Tudjarov A, Temkov D, Janevski T, Firfov O. Empirical modeling of Internet traffic at middle-level burstiness. In: *Proc. of the 12th IEEE Mediterranean on Electrotechnical Conf*. IEEE, 2004. 535–538.
- [18] Hokstad P. The G/M/m queue with finite waiting room. *Journal of Applied Probability*, 1975,12(4):779–792.
- [19] Vladimir VK. *Mathematical Methods in Queuing Theory*. Dordrecht: Kluwer Academic, 1994.
- [20] Sigman K. Appendix: A primer on heavy-tailed distributions. *Queueing Systems*, 1999,33(1/3):261–275.
- [21] WANG H. A queueing system  $GI^X/M/1/N$  with batch arrivals and randomly dropping packets mechanism. *Mathematics in Practice and Theory*, 2005,35(9):113–120 (in Chinese with English abstract).

#### 附中文参考文献:

- [21] 汪浩. 一个具有随机丢弃分组机制且分组成批到达的  $GI^X/M/1/N$  排队系统. *数学的实践与认识*, 2005,35(9):113–120.

#### 附录:定理 1 的证明

证明:i) 假设 IP 包到达路由器的时刻依次为  $\tau_0, \tau_1, \tau_2, \dots, \tau_n, \dots$ , 其中:  $\tau_0 < \tau_1 < \tau_2 < \dots < \tau_n < \dots$ , 且  $\tau_0=0$ ; 包的到达间隔为  $e_k = \tau_k - \tau_{k-1}$  ( $k=1, 2, \dots, n, \dots$ ), 且 i.i.d r.v.s  $e_k$  的分布函数为  $A(t)$ , 如图 17 所示.

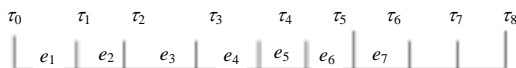


Fig.17 The interarrival times of IP packets

图 17 IP 包的到达间隔时间

AQM 算法开始丢包后,包的到达间隔变为(其中  $v_i$  为随机变量,其分布随后给出)  $e_1 + e_2 + \dots + e_{v_i}$ . 也就是说, AQM 算法开始丢包后,路由器缓存队列每隔  $e_1 + e_2 + \dots + e_{v_i}$  间隔接收一个包. 如果 i.i.d.r.v.  $e_k$  的 df  $A(t)$  是 subexponential 分布函数,根据 subexponential 分布函数的性质<sup>[20]</sup>可知:  $\Pr(e_1 + e_2 + \dots + e_{v_i} > t) \approx E[v_i] \Pr(e_1 > t)$ . 因此,  $e_1 + e_2 + \dots + e_{v_i}$  的 df  $A_i(t)$  可用如下公式估计:

$$A_i(t) = \Pr(e_1 + e_2 + \dots + e_{v_i} \leq t) = 1 - \Pr(e_1 + e_2 + \dots + e_{v_i} > t) = 1 - E[v_i][1 - A(t)].$$

由于 IP 包到达缓存队列的间隔时间变长了,因而要求  $t \geq \tau$ , 其中  $\tau$  由下式决定:

$$\int_{\tau}^{\infty} dA_i(t) = \int_{\tau}^{\infty} a_i(t) dt = \int_{\tau}^{\infty} E[v_i] a(t) dt = 1 \quad \text{即} \quad \int_{\tau}^{\infty} a(t) dt = \frac{1}{E[v_i]}.$$

如果以概率  $r_i$  按几何分布稀疏输入流(对应 AQM 按几何分布丢包),则离散型 r.v.  $v_i$  服从几何分布,从而  $E[v_i] = \frac{1}{r_i}$ ; 如果以概率  $r_i$  按均匀分布稀疏输入流(对应 AQM 按均匀分布丢包),则离散型 r.v.  $v_i$  服从均匀分布,从而,  $E[v_i] \approx \frac{1}{2} + \frac{1}{2r_i}$ .

ii) 利用 Renyi 定理(参见文献[19]第 69 页)即得.



汪浩(1962 - ),男,江西南昌人,博士生,教授,主要研究领域为拥塞控制.



严伟(1961 - ),女,副教授,主要研究领域为移动计算,网络管理.