典型 AQM 算法的性能评价模型

汪 浩[1),2)] 严 伟[1)]

[1)](北京大学信息科学技术学院网络实验室 北京 100871)

[2)](江西师范大学软件学院 南昌 330027)

摘 要 利用 GI/M/1/N 排队系统和 Internet 业务流量自相似性的特点建立了一个评价 AQM 算法在非响应业务流量下性能的分析模型,提出了利用模型的分析计算结果而不是模拟或实验手段评价 AQM 算法性能的新方法.同模拟或实验手段相比,该方法能更深刻地刻画 AQM 算法在实际网络环境中的性能.用该模型分析比较 3 个经典的 AQM 算法——TD、RED 和 GRED 的性能,所得的结果同其他研究者利用模拟或实验方法所得的结果一致.

关键词 RED 算法;主动队列管理(AQM);拥塞控制;GI/M/1/N 排队系统;重尾分布;自相似网络流量

中图法分类号 TP393

A Performance-Evaluation Model of Typical AQM Algorithms

WANG Hao[1),2)] YAN Wei[1)]

[1)](*Network Laboratory, School of Electronics Engineering and Computer Science, Peking University, Beijing* 100871)

[2)](*School of Software, Jiangxi Normal University, Nanchang* 330027)

Abstract The Internet traffic consists of responsive long-lived TCP flows and unresponsive short-lived TCP flows and UDP flows. A lot of papers have been published to model the interaction between long-lived TCP flows and AQM algorithms. However, unresponsive flows dominate the Internet traffic, which contribute 70%~80% of the Internet traffic. Therefore, it is important to model the behavior of AQM algorithms with unresponsive flows. In this paper, an analytical model is presented based on the GI/M/1/N queuing system and the self-similar traffic of the Internet to evaluate the performance of AQM algorithms with unresponsive flows. Compared with the presented experimental or simulation approaches, this modeling approach can help to reveal the underlying characteristics of the performance of AQM algorithms in a practical network environment. Using this model, the authors analyze the performance of three AQM algorithms: TD, RED and GRED.

Keywords RED algorithm; active queue management(AQM); congestion control; GI/M/1/N; heavy-tailed distributions; self-similar network traffic

1 引 言

1.1 本文的动机和贡献

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[1].UDP 流

收稿日期:2004-07-07;修改稿收到日期:2005-11-24.本课题得到国家自然科学基金(60273002)及国家"九七三"重点基础研究发展规划项目基金(2004CB318204)资助.汪 浩,男,1962 年生,博士研究生,主要研究方向为拥塞控制、QoS. E-mail: wanghao@net.pku.edu.cn; wanghao@jxnu.edu.cn.严 伟,女,1961 年生,副教授,主要研究方向为无线 Ad Hoc 网、无线 mesh 网. E-mail: yanwei@net.pku.edu.cn.

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%字节的流量属于短效 TCP 流量，55%～75%的 IP 包属于短效 TCP 流量[2,3]. 此外，2004 年度 Sigcomm 的获奖者 Simon 认为①，随着分布式多媒体应用(如视频会议、视频点播、IP 电话、远程教育)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. 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.

现有的 TCP/AQM 模型[4～8]重点放在分析 AQM 与响应业务流量相互作用的反馈机制上，但忽略了非响应业务流量对 AQM 算法性能的影响. 为了弥补 TCP/AQM 模型没有考虑非响应流的不足，已有研究者将非响应流作为“干扰信号”引入 TCP/AQM 模型，来分析非响应流对 AQM 算法性能的影响[9]. 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，所以也能利用开环的排队模型为 AQM 算法建立性能评价模型. 为此，我们利用 GI/M/1/N 排队系统和 Internet 业务流量自相似性的特点建立了一个评价 AQM 算法在非响应业务流量下性能的分析模型，提出了一个利用模型的分析计算结果而不是模拟或实验手段评价 AQM 算法性能的新方法. 该方法具有如下优越性：(1)能更真实地刻画 AQM 算法在 Internet 业务流量下的性能；(2)不需要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. 用该模型分析比较 3 个经典的 AQM 算法——TD、RED 和 GRED 的性能，所得的结果同其他研究者利用模拟或实验方法所得的结果一致[10～12].

受篇幅限制，本文重点讨论性能评价模型的相关定理及其证明和性能评价指标，该模型详细的应用我们在另文讨论.

1.2　相关研究

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[10～12]. 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[11]；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[12]. Eitan 等人通过模拟手段，分析了短效 TCP 流条件下 RED 算法的公平性[13]. 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[14].

由于非响应业务流量不能响应 AQM 算法的丢包信号，故另一些研究者试图用开环的排队模型评价 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[15]. Garetto 利用 $M^X/M/1/N$ 排队系统，分析了路由器队列长度(时延)的分布[16]；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[17].

为了用排队模型评价非响应业务流量下 AQM 算法的性能，需要知道 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，Internet 业务流量具有自相似性，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[18～20].

2　AQM 算法的性能评价模型

2.1　AQM 算法的形式化定义

为了刻画 AQM 算法的本质特征并且用排队论的方法为 AQM 算法建立性能评价模型，我们首先给 AQM 算法一个形式化的定义. 同文献[15]，我们利用瞬时队长而不是平均队长计算丢包率. 在以下的讨论中，我们用 cdf 表示“概率分布函数”，用 pdf 表示“概率密度函数”.

定义 1.　假设路由器缓存队列可容纳 N 个包，用 $m+1$ 个满足条件：

$$0=L_0<L_1<L_2<\cdots<L_m=N$$

的正整数 $L_0, L_1, L_2, \cdots, L_m$ 将缓存队列分成 m 段，每段的编号分别为 $1, 2, \cdots, m$. 如图 1 所示.

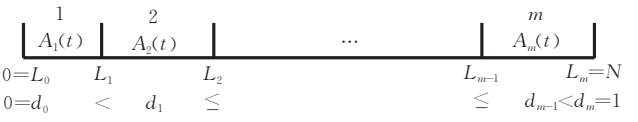

图 1　路由器缓存队列的划分

由于丢包的作用，缓存队列不同的段将具有不同的包的到达间隔(interarrival times)分布，参见图 1、图 2. 假设包到达路由器的间隔时间的 cdf 为 $A_0(t)$，如果缓存队列中当前有 k 个包，则：当 $L_0 \le k < L_1$ 时，AQM 算法以概率 $d_0=0$ 均匀丢包，此时包到达缓存队列的间隔时间的 cdf 是 $A_1(t)=A_0(t)$；当 $L_1 \le k < L_2$ 时，AQM 算法以概率 d_1 均匀丢包，此时

① Simon L.. Keynote Speech by SIGCOMM Award Winner Simon Lam. http://www.acm.org/sigs/sigcomm/sigcomm2004/conf_program.html, 2004

包到达缓存队列的间隔时间的 cdf 是 $A_2(t)$；当 $L_2 \le k < L_3$ 时，AQM 算法以概率 d_2 均匀丢包，此时包到达缓存队列的间隔时间的 cdf 是 $A_3(t)$；依次类推，当 $L_{m-1} \le k < L_m$ 时，AQM 算法以概率 d_{m-1} 均匀丢包，此时包到达缓存队列的间隔时间的 cdf 是 $A_m(t)$；当 $L_m \le k$ 时，AQM 算法以概率 $d_m = 1$ 丢包. 其中 $0 = d_0 < d_1 \le d_2 \le \cdots \le d_{m-1} < d_m = 1$，$t > 0$.

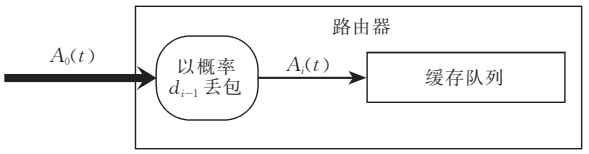

图 2 路由器丢包前后包的到达间隔分布之间的关系

对定义的几点说明：

(1)从定义中可知，当缓存队列中包的个数 $< L_1$ 时，AQM 算法不丢包；当缓存队列中包的个数 $\ge N$ 时，AQM 算法丢弃所有新到的包.

(2)$A_0(t)$ 是 AQM 算法丢包前，包到达路由器的间隔时间的 cdf，$A_1(t), A_2(t), A_3(t), \cdots, A_m(t)$ 是 AQM 算法丢包后，包到达缓存队列的间隔时间的 cdf(见图 2)；$A_1(t), A_2(t), A_3(t), \cdots, A_m(t)$ 可以通过 $A_0(t)$ 表示(具体方法见 2.2 小节定理 1)，其中 $A_1(t) = A_0(t)$.

(3)在本文的讨论中，分别用 $a_0, a_1, a_2, \cdots, a_m$ 表示 $A_0(t), A_1(t), A_2(t), A_3(t), \cdots, A_m(t)$ 的均值. 即对 $i = 0, 1, 2, \cdots, m$，

$$a_i = \int_0^\infty x \, \mathrm{d}A_i(x).$$

(4)定义中要求以概率(如 d_2)均匀丢包主要是为了避免 TCP 连接发生全局同步[21].

2.2 输入流的稀疏化

由于丢包的作用，使得到达缓存队列的 IP 包的间隔时间变长了，排队论中称之为输入流的稀疏化(thinning of recurrent flows)[22]. 以下我们讨论如何通过 $A_0(t)$ 表示 $A_i(t)$. 首先注意到，由于 $d_0 = 0$，故 $A_1(t) = A_0(t)$；对 $A_i(t)$，$i = 2, 3, \cdots, m$，我们证明了下列定理.

定理 1. 如果 $A_0(t)$ 服从 Pareto 分布，即

$$A_0(t) = 1 - \frac{1}{(1+t)^\alpha}, \quad t \ge 0, \ 1 < \alpha \le 2,$$

则 $A_i(t)$ 可用下式估计($i = 2, 3, \cdots, m$)：

$$A_i(t) = 1 - E[\nu_i] \frac{1}{(1+t)^\alpha},$$

$$t \ge (E[\nu_i])^{\frac{1}{\alpha}} - 1, \ 1 < \alpha \le 2,$$

其中离散型随机变量 ν_i 在集合 $\{1, 2, \cdots, \lceil 1/r_{i-1} \rceil\}$ 内均匀分布(这是文献[21]要求的条件)，$r_{i-1} = 1 - d_{i-1}$，且

$$Pr(\nu_i = n) = r_{i-1}, \text{ 当 } 1 \le n \le \lceil 1/r_{i-1} \rceil,$$
$$Pr(\nu_i = n) = 0, \text{ 当 } n > \lceil 1/r_{i-1} \rceil,$$
$$E[\nu_i] \approx 1/2 + 1/(2r_{i-1}).$$

证明. 根据 AQM 算法的定义，对 $i = 2, 3, \cdots, m$，AQM 算法以概率 d_{i-1} 均匀丢包，等价于以概率 $r_{i-1} = 1 - d_{i-1}$ 均匀接收包，等价于增加了 IP 包到达的间隔时间. 假设 AQM 算法丢包前，包的到达时间为 $T_0, T_1, T_2, \cdots, T_n, \cdots$，其中 $T_0 < T_1 < T_2 < \cdots < T_n < \cdots$ 且 $T_0 = 0$，包的到达间隔为

$$e_k = T_k - T_{k-1}, \ k = 1, 2, \cdots, n, \cdots,$$

e_k 为独立同分布随机变量，其 cdf 是 $A_0(t)$. AQM 算法开始丢包后，包的到达间隔为 $e_1 + e_2 + \cdots + e_{\nu_i}$，其中离散型随机变量 ν_i 在集合 $\{1, 2, \cdots, \lceil 1/r_{i-1} \rceil\}$ 内均匀分布，且

$$Pr(\nu_i = n) = r_{i-1}, \text{ 当 } 1 \le n \le \lceil 1/r_{i-1} \rceil,$$
$$Pr(\nu_i = n) = 0, \text{ 当 } n > \lceil 1/r_{i-1} \rceil,$$
$$E[\nu_i] = r_{i-1} \lceil 1/r_{i-1} \rceil (1 + \lceil 1/r_{i-1} \rceil)/2$$
$$\approx 1/2 + 1/(2r_{i-1}),$$

也就是说 AQM 算法开始丢包后，路由器缓存队列每隔 $e_1 + e_2 + \cdots + e_{\nu_i}$ 时长接收一个包，从而

$$A_i(t) = Pr(e_1 + e_2 + \cdots + e_{\nu_i} \le t)$$
$$= 1 - Pr(e_1 + e_2 + \cdots + e_{\nu_i} > t),$$

如果 IP 包到达路由器的间隔时间 e_k 满足 Pareto 分布，根据 Pareto 分布的性质，我们知道：

$$Pr(e_1 + e_2 + \cdots + e_{\nu_i} > t) \approx E[\nu_i] Pr(e_1 > t),$$

从而可用下式近似估计 $A_i(t)$：

$$A_i(t) = 1 - E[\nu_i] \frac{1}{(1+t)^\alpha},$$

$$t \ge (E[\nu_i])^{\frac{1}{\alpha}} - 1, \ 1 < \alpha < 2 \qquad \text{证毕.}$$

用类似的方法，我们可以证明如下定理 2.

定理 2. 如果 $A_0(t)$ 服从重尾 Weibull 分布，即

$$A_0(t) = 1 - \mathrm{e}^{-x^\alpha}, \ 0 < \alpha < 1,$$

则 $A_i(t)$ 可用下式估计($i = 2, 3, \cdots, m$)，其中离散型随机变量 ν_i 的含义同定理 1：

$$A_i(t) = 1 - E[\nu_i] \, \mathrm{e}^{-t^\alpha}, \ t \ge (\ln E[\nu_i])^{\frac{1}{\alpha}}.$$

根据定理 1 和定理 2，如果已知 $A_0(t)$，我们就能找出 $A_i(t)$ 的估计式($i = 2, 3, \cdots, m$)，这些 $A_i(t)$ 的估计式将用在后面的式(34)～(37)中.

2.3 AQM 性能评价模型的建立

研究人员通常通过丢包率、队列长度(含概率分布、均值和方差)、时延(时延的均值和方差)、连续丢

包数(连续丢弃 n 个包的概率分布、连续丢弃 n 个包的均值和方差)等指标评价 AQM 算法的性能. 这些指标通常可以通过模拟或实验手段获得. 实际上,这些性能指标也可以通过排队论的方法计算出来,也就是说,这些性能指标可以通过队列长度的概率分布计算出来. 因此,下一步的工作就是设法求解队列长度的概率分布. 为此,我们利用排队论中的 GI/M/1/N 模型来求解在 AQM 算法的作用下,路由器缓存队列长度的概率分布. 为简化表述,但不失一般性,我们仅对缓存队列分为两段(即 $m=2$)的情形进行讨论. 为此,先引入一些记号. 对 $t\in[0,\infty)$,用 $A_1(t)=A_0(t)$ 表示 AQM 算法丢包前(即 $d_0=0$ 时)、IP 包到达路由器缓存队列间隔时间的 cdf,其对应的 pdf 为 $a_1(t)$,均值为 a_1;当队列中包的个数 $k\geq L_1$ 时,AQM 算法以概率 d_1 均匀丢包,用 $A_2(t)$ 表示 AQM 算法丢包后(即 $d_1>0$ 时)、IP 包到达路由器缓存队列间隔时间的 cdf,其对应的 pdf 为 $a_2(t)$,均值为 a_2;路由器服务时间的 cdf 是 $B(t)$,$B(t)$ 服从参数为 μ 的(负)指数分布. 在以上假定条件下,我们证明了如下定理.

定理 3.　假设路由器缓存队列划分为两段,即 $m=2$,用 p_k 表示任意时刻路由器缓存队列中有 k 个包的概率;q_k 表示一个新包到达时,路由器缓存队列中有 k 个包的概率;a_1,a_2 分别表示 $A_1(t),A_2(t)$ 的均值;路由器服务时间服从参数为 μ 的(负)指数分布,则

$$p_k=\frac{1}{\mu a_1}q_{k-1},\ 1\leq k\leq L_1-1,$$

$$p_k=\frac{1}{\mu a_2}q_{k-1},\ L_1\leq k\leq N,$$

且
$$p_0=1-\sum_{k=1}^{N}p_k.$$

证明.　假设 $Q(t)$ 表示 t 时刻路由器缓存队列中包的数目,$U(t)$ 表示从 t 时刻起至下一个包到达时刻的时间(residual life). 路由器服务时间的 cdf 是 $B(t)$,$B(t)$ 服从参数为 μ 的(负)指数分布. 二元组$(Q(t),U(t))$构成连续时间离散状态的马氏链. 在上述假定下,为了证明定理,我们利用辅助变量法(supplementary variables)[23,24] 建立 Kolmogorov-Chapman 前向方程. 记

$$p_k(x,t)\mathrm{d}x\triangleq Pr(Q(t)=k,\ x<U(t)\leq x+\mathrm{d}x).$$

根据 AQM 算法定义中的概率特征,马氏过程在 t 和 $t+\Delta$ 时刻的状态有如下关系:

$$p_0(x-\Delta,t+\Delta)=p_0(x,t)+p_1(x,t)\mu\Delta+o(\Delta) \tag{1}$$

$$p_k(x-\Delta,t+\Delta)=p_k(x,t)(1-\mu\Delta)+p_{k+1}(x,t)\mu\Delta+a_1(x)\int_0^{\Delta}p_{k-1}(w,t)\mathrm{d}w+o(\Delta),\ 1\leq k\leq L_1 \tag{2}$$

$$p_k(x-\Delta,t+\Delta)=p_k(x,t)(1-\mu\Delta)+p_{k+1}(x,t)\mu\Delta+a_2(x)\int_0^{\Delta}p_{k-1}(w,t)\mathrm{d}w+o(\Delta),\ L_1+1\leq k\leq N-1 \tag{3}$$

$$p_N(x-\Delta,t+\Delta)=p_N(x,t)(1-\mu\Delta)+a_2(x)\Big[\int_0^{\Delta}p_{N-1}(w,t)\mathrm{d}w+\int_0^{\Delta}p_N(w,t)\mathrm{d}w\Big]+o(\Delta) \tag{4}$$

将式(1)～(4)变化为

$$\frac{p_0(x-\Delta,t+\Delta)-p_0(x,t)}{\Delta}=\mu p_1(x,t)+\frac{o(\Delta)}{\Delta},$$

$$\frac{p_k(x-\Delta,t+\Delta)-p_k(x,t)}{\Delta}=-\mu p_k(x,t)+\mu p_{k+1}(x,t)+\frac{a_1(x)\int_0^{\Delta}p_{k-1}(w,t)\mathrm{d}w}{\Delta}+\frac{o(\Delta)}{\Delta},\ 1\leq k\leq L_1,$$

$$\frac{p_k(x-\Delta,t+\Delta)-p_k(x,t)}{\Delta}=-\mu p_k(x,t)+\mu p_{k+1}(x,t)+\frac{a_2(x)\int_0^{\Delta}p_{k-1}(w,t)\mathrm{d}w}{\Delta}+\frac{o(\Delta)}{\Delta},\ L_1+1\leq k\leq N-1,$$

$$\frac{p_N(x-\Delta,t+\Delta)-p_N(x,t)}{\Delta}=-\mu p_N(x,t)+\frac{a_2(x)\Big[\int_0^{\Delta}p_{N-1}(w,t)\mathrm{d}w+\int_0^{\Delta}p_N(w,t)\mathrm{d}w\Big]}{\Delta}+\frac{o(\Delta)}{\Delta}.$$

对等式两边当 $\Delta\to 0$ 时取极限,得

$$\frac{\partial p_0(x,t)}{\partial t}-\frac{\partial p_0(x,t)}{\partial x}=\mu p_1(x,t) \tag{5}$$

$$\frac{\partial p_k(x,t)}{\partial t}-\frac{\partial p_k(x,t)}{\partial x}=-\mu p_k(x,t)+\mu p_{k+1}(x,t)+a_1(x)p_{k-1}(0,t),\ 1\leq k\leq L_1 \tag{6}$$

$$\frac{\partial p_k(x,t)}{\partial t}-\frac{\partial p_k(x,t)}{\partial x}=-\mu p_k(x,t)+\mu p_{k+1}(x,t)+a_2(x)p_{k-1}(0,t),\ L_1+1\leq k\leq N-1 \tag{7}$$

$$\frac{\partial p_N(x,t)}{\partial t}-\frac{\partial p_N(x,t)}{\partial x}=-\mu p_N(x,t)+a_2(x)[p_{N-1}(0,t)+p_N(0,t)] \tag{8}$$

我们考虑当 $t\to\infty$ 时马氏链处于稳定状态时的情

况,记

$$p_k(x)=\lim_{t\to\infty}p_k(x,t),\ 0\le k\le N \tag{9}$$

此时$\frac{\partial p_k(x)}{\partial t}=0$,故方程式(5)~(8)简化为

$$-\frac{\mathrm{d}p_0(x)}{\mathrm{d}x}=\mu p_1(x) \tag{10}$$

$$-\frac{\mathrm{d}p_k(x)}{\mathrm{d}x}=-\mu p_k(x)+\mu p_{k+1}(x)+a_1(x)p_{k-1}(0),\quad 1\le k\le L_1 \tag{11}$$

$$-\frac{\mathrm{d}p_k(x)}{\mathrm{d}x}=-\mu p_k(x)+\mu p_{k+1}(x)+a_2(x)p_{k-1}(0),\quad L_1+1\le k\le N-1 \tag{12}$$

$$-\frac{\mathrm{d}p_N(x)}{\mathrm{d}x}=-\mu p_N(x)+a_2(x)[p_{N-1}(0)+p_N(0)] \tag{13}$$

式(10)~(13)共计 $N+1$ 个微分方程,刻画了路由器缓存队列在稳定状态时队列中有 k 个包的概率.引入记号:

$$p_k^*(s)\triangleq\int_0^\infty \mathrm{e}^{-sx}p_k(x)\mathrm{d}x,\ 0\le k\le N \tag{14}$$

$$a_1^*(s)\triangleq\int_0^\infty \mathrm{e}^{-st}\mathrm{d}A_1(t) \tag{15}$$

$$a_2^*(s)\triangleq\int_0^\infty \mathrm{e}^{-st}\mathrm{d}A_2(t) \tag{16}$$

$$p_k\triangleq p_k^*(0)=\int_0^\infty p_k(x)\mathrm{d}x,\ 0\le k\le N \tag{17}$$

$$a_1\triangleq\int_0^\infty x\,\mathrm{d}A_1(x)=-\left.\frac{\mathrm{d}a_1^*(s)}{\mathrm{d}s}\right|_{s=0} \tag{18}$$

$$a_2\triangleq\int_0^\infty x\,\mathrm{d}A_2(x)=-\left.\frac{\mathrm{d}a_2^*(s)}{\mathrm{d}s}\right|_{s=0} \tag{19}$$

$$q_k\triangleq a_1p_k(0),\ 0\le k\le L_1-1 \tag{20}$$

$$q_k\triangleq a_2p_k(0),\ L_1\le k\le N \tag{21}$$

其中,式(14)~(16)是相关函数的 Laplace-Stieltjes(简称 L-S)变换,p_k 是任意时刻路由器缓存队列中有 k 个包的概率;q_k 是一个新包到达时,路由器缓存队列中有 k 个包的概率.对式(10)~(13)做 L-S 变换,并整理后得

$$sp_0^*(s)=-\mu p_1^*(s)+p_0(0) \tag{22}$$

$$(s-\mu)p_k^*(s)=-\mu p_{k+1}^*(s)-a_1^*(s)p_{k-1}(0)+p_k(0),\quad 1\le k\le L_1 \tag{23}$$

$$(s-\mu)p_k^*(s)=-\mu p_{k+1}^*(s)-a_2^*(s)p_{k-1}(0)+p_k(0),\quad L_1+1\le k\le N-1 \tag{24}$$

$$(s-\mu)p_N^*(s)=-a_2^*(s)[p_{N-1}(0)+p_N(0)]+p_N(0) \tag{25}$$

在式(22)~(25)中令 $s=0$,并注意到 $a_1^*(0)=1$,$a_2^*(0)=1$,$p_k\triangleq p_k^*(0)$,经化简后得

$$\mu p_1=p_0(0) \tag{26}$$

$$\mu(p_k-p_{k+1})=p_{k-1}(0)-p_k(0),\ 1\le k\le N-1 \tag{27}$$

$$\mu p_N=p_{N-1}(0) \tag{28}$$

由式(26)~(28)并利用数学归纳法知

$$\mu p_k=p_{k-1}(0),\ 0\le k\le N \tag{29}$$

由式(20),(21)和(29)我们得到定理 2 的结论:

$$p_k=\frac{1}{\mu a_1}q_{k-1},\ 1\le k\le L_1-1 \tag{30}$$

$$p_k=\frac{1}{\mu a_2}q_{k-1},\ L_1\le k\le N \tag{31}$$

且

$$p_0=1-\sum_{k=1}^{N}p_k.$$

证毕.

定理 2 可以推广到 $m>2$ 的情况,从而有下面定理.

定理 4. 假设路由器缓存队列划分为 m 段,且 $m>2$,用 p_k 表示任意时刻路由器缓存队列中有 k 个包的概率;q_k 表示一个新包到达时,路由器缓存队列中有 k 个包的概率;$a_1,a_2,\cdots,a_m$ 分别表示 $A_1(t)$,$A_2(t),A_3(t),\cdots,A_m(t)$ 的均值;路由器服务时间服从参数为 μ 的(负)指数分布,则

$$p_k=\frac{1}{\mu a_1}q_{k-1},\ 1\le k\le L_1-1,$$

$$p_k=\frac{1}{\mu a_2}q_{k-1},\ L_1\le k\le L_2-1,$$

$$p_k=\frac{1}{\mu a_3}q_{k-1},\ L_2\le k\le L_3-1,$$

$$\vdots$$

$$p_k=\frac{1}{\mu a_m}q_{k-1},\ L_{m-1}\le k\le L_m=N$$

且

$$p_0=1-\sum_{k=1}^{N}p_k.$$

根据定理 3 或定理 4,如果已知 $q_k(k=0,1,\cdots,N)$,我们能方便地求出 $p_k(k=0,1,\cdots,N)$.以下我们采用文献[23]的方法,利用 GI/M/1/N 排队模型的嵌入式马氏链求 $q_k(k=0,1,\cdots,N)$.q_k 可以通过求解以下线性方程组获得

$$q_j=\sum_{i=0}^{N}q_ia_{i,j},\ j=0,1,\cdots,N \tag{32}$$

$$\sum_{j=0}^{N}q_j=1 \tag{33}$$

其中 $a_{i,j}$ 的计算公式如下所示:

$$a_{i,j}=\int_0^{\infty}\frac{\mathrm{e}^{-\mu t}(\mu t)^{i+1-j}}{(i+1-j)!}\mathrm{d}A_1(x),$$

$$0\le i\le L_1-1,\ 1\le j\le i+1 \tag{34}$$

$$a_{i,j}=\int_0^{\infty}\frac{\mathrm{e}^{-\mu t}(\mu t)^{i+1-j}}{(i+1-j)!}\mathrm{d}A_2(x),$$

$$L_1\le i\le L_2-1,\ 1\le j\le i+1 \tag{35}$$

$$a_{i,j}=\int_0^{\infty}\frac{\mathrm{e}^{-\mu t}(\mu t)^{i+1-j}}{(i+1-j)!}\mathrm{d}A_3(x),$$

$$L_2\le i\le L_3-1,\ 1\le j\le i+1 \tag{36}$$

$$\vdots$$

$$a_{i,j}=\int_0^{\infty}\frac{\mathrm{e}^{-\mu t}(\mu t)^{i+1-j}}{(i+1-j)!}\mathrm{d}A_m(x),$$

$$L_{m-1}\le i\le L_m-1=N-1,\ 1\le j\le i+1 \tag{37}$$

$$a_{N,j}=a_{N-1,j},\ 0\le j\le N \tag{38}$$

$$a_{i,0}=1-\sum_{j=1}^{i+1}a_{i,j},\ 0\le i\le N \tag{39}$$

求得 $q_k(k=0,1,\cdots,N)$ 之后，可以利用定理 3 或定理 4 计算 $p_k(k=0,1,\cdots,N)$.

3　评价 AQM 算法性能的指标体系

我们将通过丢包率、队列长度(含概率分布、均值和方差)、时延(时延的均值和方差)、连续丢包数(连续丢弃 n 个包的概率分布、连续丢弃 n 个包的均值和方差)等指标评价 AQM 算法的性能. 利用 2.3 小节求出的队列长度的概率分布 p_k 和 q_k，可以计算上述各指标. 以下是各评价指标的具体计算公式.

3.1　丢包率的计算公式

(1) TD 算法的丢包率：当缓存队列满时，TD 算法开始丢包，故

$$\text{TD 算法的丢包率}=q_N \tag{40}$$

(2) 其它 AQM 算法的丢包率：当缓存队列包的个数超过 L_1 时开始丢包，故

$$\text{其它 AQM 算法的丢包率}=\sum_{i=L_1}^{N}q_i d_{u(i)} \tag{41}$$

其中函数 $u(i)$ 定义为

$$u(i)=\begin{cases}0, & L_0\le i<L_1\\ 1, & L_1\le i<L_2\\ 2, & L_2\le i<L_3\\ \vdots & \\ m-1, & L_{m-1}\le i\le N\end{cases}.$$

3.2　与队列长度有关的指标的计算公式

用随机变量 ζ 表示任意时刻路由器缓存队列的长度，用随机变量 η 表示一个新包到达时，路由器缓存队列的长度，则

$$Pr(\zeta=k)=p_k \tag{42}$$

$$\bar{\zeta}\triangleq E[\zeta]=\sum_{k=1}^{N}kp_k \tag{43}$$

$$Var[\zeta]=\sum_{k=1}^{N}(k-\bar{\zeta})^2p_k \tag{44}$$

$$Pr(\eta=k)=q_k \tag{45}$$

$$\bar{\eta}\triangleq E[\eta]=\sum_{k=1}^{N}kq_k \tag{46}$$

$$Var[\eta]=\sum_{k=1}^{N}(k-\bar{\eta})^2q_k \tag{47}$$

3.3　与时延有关的指标的计算公式

用随机变量 ξ 表示 IP 包在路由器中的等待时间，则

$$\rho\triangleq Pr(\text{包需要在路由器缓存队列等待})=\sum_{k=1}^{N}q_k \tag{48}$$

$$\bar{\xi}\triangleq E[\xi\mid\text{包需要在路由器缓存队列等待}]=\frac{1}{\rho}\sum_{k=1}^{N}q_k\frac{k}{\mu} \tag{49}$$

$$Var[\xi\mid\text{包需要在路由器缓存队列等待}]=\frac{1}{\rho}\sum_{k=1}^{N}\left(\frac{k}{\mu}-\bar{\xi}\right)^2q_k \tag{50}$$

3.4　与连续丢包数有关的指标的计算公式

连续丢包数是衡量一个 AQM 算法是否会引起 TCP 连接全局同步的重要指标.

3.4.1　TD 连续丢包数的计算公式

由于 IP 包到达路由器的间隔时间构成一个更新过程，记该更新过程剩余时间的分布为 $V(x)$，由更新定理[25]知 $V(x)$ 可近似估计为

$$V(x)=\frac{1}{a_0}\int_0^x(1-A_0(t))\mathrm{d}t=1-(1+x)^{1-\alpha}.$$

如果缓存队列已满，则下一个包被丢弃的概率 p 为

$$p=\int_0^{\infty}V(x)\mathrm{d}B(x)=\int_0^{\infty}\mu V(x)\mathrm{e}^{-\mu x}\mathrm{d}x \tag{51}$$

假设用随机变量 γ_{TD} 表示 TD 算法的连续丢包数，则[15]

$$Pr(\gamma_{\mathrm{TD}}>n)=p^n,\ \forall n\ge 0 \tag{52}$$

$$\bar{\gamma}_{\mathrm{TD}}\triangleq E(\gamma_{\mathrm{TD}})=\sum_{n=1}^{\infty}Pr(\gamma_{\mathrm{TD}}\ge n)=\frac{1}{1-p} \tag{53}$$

$$Var(\gamma_{TD})=\sum_{n=1}^{\infty}(2n-1)Pr(\gamma_{TD}\geq n)-\bar{\gamma}_{TD}^2$$
$$=\frac{p}{(1-p)^2} \quad (54)$$

3.4.2 其它 AQM 连续丢包数的计算公式

假设用随机变量 γ_{AQM} 表示其它 AQM 算法的连续丢包数,则[15]

$$Pr(\gamma_{AQM}>n)=\frac{\sum_{i=L_1}^{N}q_i d_{u(i)}^{n+1}}{\sum_{i=L_1}^{N}q_i d_{u(i)}},\ \forall n\geq 0 \quad (55)$$

$$\bar{\gamma}_{AQM}\triangleq E(\gamma_{AQM})=\sum_{n=1}^{\infty}Pr(\gamma_{AQM}\geq n)$$
$$=\frac{\sum_{i=L_1}^{N}q_i\frac{d_{u(i)}}{1-d_{u(i)}}}{\sum_{i=L_1}^{N}q_i d_{u(i)}} \quad (56)$$

$$Var(\gamma_{AQM})=\sum_{n=1}^{\infty}(2n-1)Pr(\gamma_{AQM}\geq n)-\bar{\gamma}_{AQM}^2$$
$$=\frac{\sum_{i=L_1}^{N}q_i\frac{(1+d_{u(i)})d_{u(i)}}{[1-d_{u(i)}]^2}}{\sum_{i=L_1}^{N}q_i d_{u(i)}}-\bar{\gamma}_{AQM}^2 \quad (57)$$

3.5 链路利用率的计算公式

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,

AQM 算法的链路利用率 $=1-p_0$ (58)

4 AQM 算法 TD、RED 和 GRED 的性能评价

本小节我们将利用 AQM 性能评价模型及其指标体系评价 3 个具体的 AQM 算法 TD、RED 和 GRED. 限于篇幅,我们在此仅给出主要的结论,更详细的分析评价结果见文献[26].

根据第 3 节的性能评价指标体系,并按表 1 的参数设置,我们分别对 $N=30$,$N=40$ 和 $N=50$ 计算了 TD、RED 和 GRED 算法在不同的输入强度下的性能指标,其中,输入强度 $=(\alpha-1)/\mu$,α 和 μ 分别是 Pareto 和指数服务时间的参数. 计算程序用 Maple 编写,TD、RED 和 GRED 算法的参数设置见表 1. 通过对计算结果的分析,我们发现,TD、RED 和 GRED 算法的差异主要体现在丢包率、连续丢包数的均值、连续丢包数的方差等评价指标上,其它评价指标几乎看不出区别;此外,RED 和 GRED 算法的各项评价指标也差别不大. 限于篇幅,我们仅给出 $N=50$ 的计算结果.

表 1 TD、RED 和 GRED 算法的参数设置

	$N=30$			$N=40$			$N=50$		
	max_p	min_{th}	max_{th}	max_p	min_{th}	max_{th}	max_p	min_{th}	max_{th}
TD 算法	NA	NA	NA	NA	NA	NA	NA	NA	NA
RED 算法	0.1	10	30	0.1	10	40	0.1	15	50
GRED 算法	0.1	10	20	0.1	10	30	0.1	15	35

(1)从图 3 可以看出,在不同的输入强度下,RED 算法和 GRED 算法的丢包率高于 TD 算法,参见式(40),(41),这一结果也被许多模拟和实验结果所证实,例如文献[10,12]. 也就是说,RED 算法或 GRED 算法是通过提高丢包率来避免 TCP 连接的全局同步.

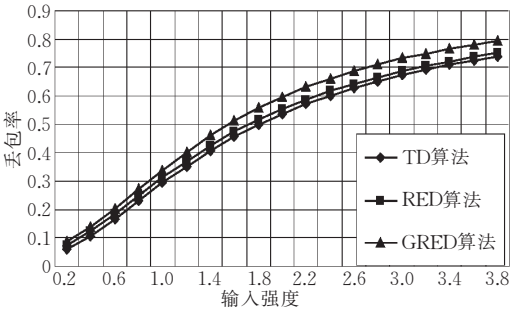

图 3 丢包率

(2)从图 4,图 5 可以看出,RED 算法连续丢包数的均值、连续丢包数的方差优于 GRED 算法和 TD 算法,参见式(53)~(57),这表明 RED 算法在发生拥塞时,能有效避免 TCP 连接全局同步的问题,其中 RED 算法连续丢包数的均值接近于 0,且方差接近于 0;GRED 算法连续丢包数的均值小于 1.5,且方差接近于 1. 因此 RED 算法比 GRED 算法还要好.

(3)在不同的输入强度下,TD、RED 和 GRED 算法其它的性能评价指标几乎看不出区别,如包需要在路由器缓存队列等待的概率(式(48))、新包到达时缓存队列中包的个数的均值(式(46))、新包到达时缓存队列中包的个数的方差(式(47))、任意时

刻缓存队列中包的个数的均值(式(43))、任意时刻缓存队列中包的个数的方差(式(44))、平均时延(式(49))、时延的方差即抖动(式(50)). 这些结果与文献[10～12]的实验结论一致. 限于篇幅,其相应图形省略.

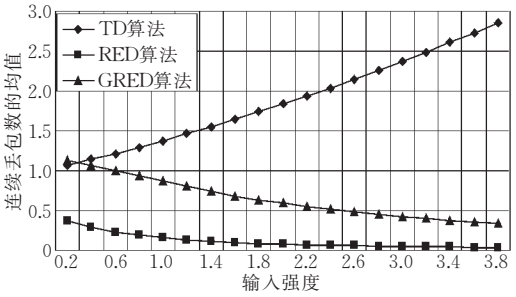

图 4　连续丢包率的均值

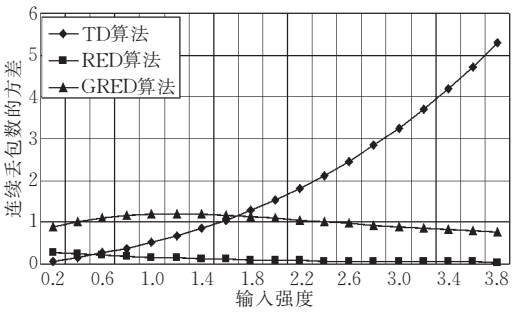

图 5　连续丢包率的方差

5　总结和进一步的研究内容

本文利用 GI/M/1/N 排队系统和 Internet 业务流量自相似性的特点建立了一个评价 AQM 算法在非响应业务流量下性能的分析模型,提出了利用模型的分析计算结果而不是模拟或实验手段评价 AQM 算法性能的新方法. 同模拟或实验手段相比,该方法能更深刻地刻画 AQM 算法在实际网络环境中的性能. 用该模型分析比较 3 个经典的 AQM 算法(TD、RED 和 GRED)的性能,所得的结果同其他研究者利用模拟或实验方法所得的结果一致.

参　考　文　献

1　Youngmi J., Vinay R., Anja F.. On the impact of variability on the buffer dynamics in IP networks. In: Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing, Allerton, 1999, 1～13

2　Brownlee N.. Understanding Internet traffic streams: Dragonflies and tortoises. IEEE Communications Magazine, 2002, 40(10): 110～117

3　Thompson K., Miller G. J. Wilder R.. Wide-area Internet traffic patterns and characteristics. IEEE Network, 1997, 11(6): 10～23

4　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

5　Mario B., Alfio L., Giovanni A.. A fluid-based model of time-limited TCP flows. Computer Networks, 2004, 44(3): 275～288

6　Tinnakornsrisuphap P., Richard J. La.. Characterization of queue fluctuations in probabilistic AQM mechanisms. ACM SIGMETRICS Performance Evaluation Review, 2004, 32(1): 283～294

7　Jae C., Mark C.. Analysis of active queue management. In: Proceedings of the 2nd IEEE International Symposium on Network Computing and Applications, Massachusetts, 2003, 359～366

8　Alazemi H. M., Mokhtar A., Azizoglu M.. Stochastic approach for modeling random early detection gateways in TCP/IP networks. In: Proceedings of the IEEE ICC 2001, Helsinki, 2001, 2385～2390

9　Hollot C. V., Liu Y., Misra V. *et al*. Unresponsive flows and AQM performance. In: Proceedings of the IEEE INFOCOM 2003, San Francisco, 2003, 85～95

10　Trinh A. T., Molnar S. A.. Comprehensive performance analysis of random early detection mechanism. Telecommunication System, 2004, 25(1,2): 9～31

11　Brandauer C., Iannaccone G., Diot C. *et al*. Comparison of tail drop and active queue management performance for bulk-data and web-like Internet traffic. In: Proceedings of the 6th IEEE Symposium on Computers and Communications, Tunisia, 2001, 3～5

12　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

13　Eitan A., Tania J.. Simulation analysis of RED with short lived TCP connections. Computer Network, 2004, 44(5): 631～641

14　Christiansen M., Jeffay K., Ott D. *et al*. Tuning RED for web traffic. IEEE/ACM Transactions on Networking, 2001, 9(3): 249～269

15　Bonald T.. Analytic evaluation of RED performance. In: Proceedings of the IEEE INFOCOM 2000, Israel, 2000, 1415～1424

16　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

17　Younsuk K., Kiseon K.. Loss probability behavior of Pareto/M/1/K queue. IEEE Communication Letters, 2003, 7(1):

39～41

18 Cao J., Cleveland W. S., Dong L. *et al*. Internet traffic tends toward Poisson and independent as the load increase. In: Denison D. D., Hansen M. H., Holmes C. C. *et al*. eds.. Nonlinear Estimation and Classification. New York: Springer, 2002, 83～109

19 Cao J., Cleveland W. S., Dong L. *et al*. On the nonstationarity of Internet traffic. ACM SIGMETRICS Performance Evaluation Review, 2001, 29(1): 102～112

20 Tudjarov A., Temkov D., Janevski T. *et al*. Empirical modeling of Internet traffic at middle-level burstiness. In: Proceedings of the 12th IEEE Electrotechnical Conference, Mediterranean, 2004, 535～538

21 Floyd S., Jacobson V.. Random early detection gateways for congestion avoidance. IEEE/ACM Transactions on Networking, 1993, 1(4): 397～413

22 Vladimir V. K.. Mathematical Methods in Queuing Theory. Dordrecht: Kluwer Academic, 1994

23 Vijaya P. L., Gupta U. C.. On the finite-buffer bulk-service queue with general independent arrivals: GI/M$^{[b]}$/1/N. Operations Research Letters. 1999, 25(5): 241～245

24 Hokstad P.. The G/M/m queue with finite waiting room. Journal of Applied Probability, 1975, 12(4): 779～792

25 Samuel K., Howard M. T.. A First Course in Stochastic Processes. Boston: Academic Press, 1975, 192～197

26 Wang H., Yan W., Huang M. H.. An analytic model for evaluating the performance of AQM algorithms with self-similar traffic. In: Proceedings of the 11th Joint International Computer Conference, Chongqing, China, 2005, 1～6

WANG Hao, born in 1962, Ph. D. candidate. His research interests include QoS, congestion control.

YAN Wei, born in 1961, associate professor. Her research interests include wireless Ad Hoc network, wireless mesh network.

Background

AQM(Active Queue Management) is an important approach for congestion control in Internet, which has drawn significant attention from researchers. In order to characterize the behavior of AQM, it is necessary to develop a mathematical model for its behavior. The Internet traffic consists of responsive long-lived TCP flows and unresponsive short-lived TCP flows and UDP flows. A lot of papers have been published to model the interaction between long-lived TCP flows and AQM algorithms based on closed-loop control system. However, short lived TCP flows and UDP flows dominate the Internet and account for about 70%～80% of the traffic. Therefore, it is important to model the behavior of AQM algorithms with unresponsive flows.

Because the short lived TCP flows and UDP flows are unresponsive to the AQM's dropping signal, an open queuing system is used to establish the performance evaluation model for AQM algorithms with these flows. Hence, the authors established three extended queuing systems by embedding AQM into the original GI/M/1/N and GI^X/M/1/N queuing systems. These three extended queuing systems are a queuing system GI/M/1/N with thinning of input flows, a queuing system GI^X/M/1/N with thinning of input flows, and a queuing system GI^X/M/1/N with balking respectively. In this paper, an analytical model is presented based on the extended GI/M/1/N queuing system with thinning of input flows and the self-similar traffic of Internet to evaluate the performance of AQM algorithms with unresponsive flows.