

基于时序推断的拥塞控制策略的性能分析

许立波 吴国新

(东南大学计算机网络与信息集成重点实验室 南京 210096)

摘 要 通过对 TD、RED 等算法进行较精确的 Markov 建模,以理论分析和数值模拟的方法对各种拥塞控制策略的性能进行深入的研究和比较,并着重从时间序列的角度分析平均队列长度计算方法的数学过程和理论依据,进而提出基于时序预测推断的平均队列长度的计算方法,给出相应的数学模型和数值模拟性能曲线,结果证明:新的算法在系统稳定性和资源利用率上表现出明显的优越性,其走势更加接近理想的性能曲线.最后考察了各算法结合优先级丢弃策略的性能表现.

关键词 拥塞控制;平均队列长度;时间序列;马尔可夫链;EWMA

中图法分类号 TP393

Analysis on Congestion Control Strategy Based on Time Series Deduction

XU Li-Bo WU Guo-Xin

(Key Laboratory of Computer Network and Information Integration, Southeast University, Nanjing 210096)

Abstract Congestion control is one of key technologies in high-speed networks. Through building accurate Markov chain network models for TD and RED arithmetics, the performances of various congestion control strategies are deeply studied and compared by theoretical analysis and numerical simulation. In particular the mathematics process and theoretic source of average queue size are analyzed from time series measures. Then a new calculating formula based on time series deduction and its network model are given. The results show that new arithmetic brings obvious advantages on reliability and resource utilization. Its trend is more close to ideal value curve. At last the performances of arithmetics with various priorities discarding strategy are reviewed.

Keywords congestion control; average queue size; time series; Markov chain; EWMA

1 引 言

目前的 Internet 采用端到端的拥塞控制策略来避免拥塞塌陷(congestion collapse)的发生^[1].端系统(如 TCP)将包丢失看作一种隐式拥塞信号并随之降低发送速率;中继系统(如路由器)则设置平均队列长度以容纳流的突发性,构建丢弃率函数以分级化丢包过程.尽管尾丢弃策略简单易用,但它在流突发期间的连续丢包导致了网络的不稳定及对突发源的偏见,因此 IETF 建议采用类似 RED 算法的

AQM 的形式来设计链路拥塞控制策略^[2],而 RED 已经应用于一些商业路由器. RED 的主要特征有两个:丢弃率函数和平均队列长度. RED 的基本原理是将平均队列长度看作拥塞发生的信号,并将其通过线性变换(也称丢弃率函数)映射成丢弃率. RED 设置两个门限值,当平均队列长度处于门限值之间时,按照丢弃率函数执行丢包行为,当平均队列长度高于高门限值时,则实行全丢弃.

由于在 Internet 中 TCP 流占了较大比重,类 TCP 流的 AIMD(Additive-Increase Multiplicative Decrease)行为使得 RED 的门限值和平均队列长度

变得有意义,也使得 RED 可以通过一定数量的丢包行为有效地控制网络流量。但在组播应用中,为避免反馈风暴和保证链路高可用,大多使用非响应性(unresponsive)流,它的 aggressive 行为严重影响了 RED 的性能以及同一链路中响应性(responsive)流的运作,因此怎样使非响应性流具有 AIMD 的特征成为组播拥塞控制研究的热点,并发展出基于窗口(如 RLA、PGMCC)和基于等式(如 LTRC、TFMCC)的两种控制方案。一些 RED 的变种(如 ZL-RED^[3])部分解决了抑制竞争流的问题。

通常,对 RED 算法的各种性能测试及参数估计大多通过仿真或测量取得,因为 RED 中的参数与网络流的 AIMD 性质形成复杂的互耦关系且难于显式建模表达,使得算法缺少理论推导过程和理论依据。文献[4]是较早的通过概率分析对 ATM 的 RED 框架简略建模的理论分析文献,后有学者将其扩展到 TCP/IP 网络,通过二阶马尔可夫链对丢弃率函数的作用进行较精确建模,并考虑了 ECN 反馈的影响^[5]。文献[6]是考察 RED 综合性能的较全面的文献,也采用随机过程类的建模方法。但这些分析模型都使用即时队列来近似平均队列长度,因此对后者的性能影响缺乏足够重视。文献[7]采用流体模型,部分考虑了平均队列长度的作用,但它关注系统的瞬态行为而非稳态行为。文献[8-9]以具体的 TCP 版本为背景,以 TCP 拥塞窗口为基础,结合 RED 考察网络的各项性能,这类方法灵活精确,但大多适用于单一类型流的分析,不适用于聚合流或大规模流的分析。本文在对传统各种拥塞控制策略进行建模和分析的同时,着重分析平均队列长度的性能影响和数学表达,考察系统队列长时间尺度上的稳态行为表现,并在此基础上提出了基于时序推断的参数建模方法,通过数值模拟验证了新的研究策略的价值。

2 系统模型定义及推导

为便于讨论,我们采用文献[5]的定义,假设系统具有如下的特征:缓存队列的长度有限;包到达服从一定参数的泊松分布;包长度是固定的,因此每个包的发送所需时间固定,设为 Δt ;包丢弃决策按照每 Δt 执行一次的频率,且一旦决定丢弃则在 Δt 内到达的包都被丢弃。在此前提下考虑 Δt 期间发生的变化。以下是建模过程所需的变量定义:

$X(k)$: 在 t_k 时刻的队列长度。

$Y(k)$: 在 $\Delta t(t_{k+1} - t_k)$ 时间内到达的包数量。

B : 队列的大小。

$Z(k)$: 一个二值的布尔变量,有

$Z(k)=1$ 表示在 $\Delta t(t_{k+1} - t_k)$ 内到达的 $Y(k)$ 个包被放入队列;

$Z(k)=0$ 表示在 $\Delta t(t_{k+1} - t_k)$ 内到达的 $Y(k)$ 个包被丢弃。

$F(n)$: 丢弃率函数,有

$$F(n) = P[Z(k) = 0 \mid X(k) = n], n \geq 0;$$

$$F(n) \in [0, 1];$$

$$F(0) = 0, F(B) = 1;$$

在 RED 中则有

$$F(n) = \begin{cases} 0, & n \leq \min; \\ \frac{p_{\max}(n - \min)}{\max - \min}, & \min < n < \max; \\ 1, & n \geq \max. \end{cases}$$

其中, \min 和 \max 为设定的门限值。

$$\bar{F}(n): \bar{F}(n) = 1 - F(n).$$

λ_{id} : 没有包丢失情况下的泊松到达参数,包含三个重复 ACK 引起的窗口减半情况。

λ_{to} : 包丢失情况下的泊松到达参数。

$U(x)$: 一个二值的布尔变量,并有如下表达:

$$U(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}.$$

定理 1. 在上述条件下, $X(k) (k=0, 1, 2, \dots)$ 可近似看作一个有限状态的二阶 Markov 链。

易见 $X(k)$ 是一个随机过程(嵌入的 Markov 链),考虑 $k+1$ 时刻的队列中包数量,有

$$X(k+1) = X(k) - U(X(k)) + Y(k)Z(k).$$

我们需要证明 $X(k+1)$ 只与 $X(k)$ 及 $X(k-1)$ 相关。设 f 为任意映射算子,先考察 $Z(k)$ 及其定义,显然 $Y(k)$ 个包是否被丢弃取决于 t_k 时刻的队列长度 $X(k)$,因此 $Z(k) = f(X(k))$;再考察 $Y(k)$ 及其定义,由于 AIMD 行为的存在, $\Delta t(t_{k+1} - t_k)$ 时间内到达的包数量受到上一时段 $\Delta t(t_k - t_{k-1})$ 内是否曾执行过包丢弃决定的影响,即 $Z(k-1)$ 的值,因此 $Y(k) = f(Z(k-1))$,又由于 $Z(k) = f(X(k))$,故 $Y(k) = f(Z(k-1)) = f(f(X(k-1)))$,这样我们得到

$$X(k+1) = X(k) - U(X(k)) + [f(f(X(k-1)))] [f(X(k))],$$

即 $X(k+1) = f[X(k), X(k-1)]$,又 $0 \leq k \leq \max$, $X(k)$ 有限,故所求得证。

定理 2. 低丢失率下,队列在拥塞避免阶段的泊松到达参数期望值近似为 $\lambda_{\text{id}}(1-p)$,其中 p 是包丢失率。

有 $E[\lambda] = \lambda_{\text{id}}(1-p) + \lambda_{\text{to}}p$,考查在长期稳定

状态下 λ_{to} 与 λ_{td} 的数值关系:根据文献[10]的推证,在 TD(Triple-Duplicate)期间包输出量 $E[Y_{td}] = \frac{E[W]}{3} \left(\frac{1-p}{p} + E[W] \right)$,在 TO(Time-Outs)期间包

输出量 $E[Y_{to}] = \frac{1}{1-p}$,则

$$\lambda_{to} = \lambda_{td} \frac{E[Y_{to}]}{E[Y_{td}]} \approx \lambda_{td} \frac{1}{\frac{1-p}{3p} \sqrt{\frac{4}{3p} + \frac{4}{9p}}} \\ \approx \lambda_{td} \lim_{p \rightarrow 0} \frac{1}{\frac{1-p}{3p} \sqrt{\frac{4}{3p} + \frac{4}{9p}}} = 0,$$

故得

$$E[\lambda] = \lambda_{td}(1-p) = \lambda_{td}\bar{p}.$$

记拥塞门限值 min, max 分别为 $n, n+r$. 对于非响应性流,该随机过程 $X(k)$ 是一个带二级门限值的可变输入率的排队模型,设系统平稳状态后的极限概率为 $\pi_i (i=0, 1, 2, \dots, n+r)$,按 K 式方程可以得到

$$\begin{cases} \lambda \pi_0 = \pi_1, \lambda^2 \pi_0 = \pi_2, \dots, \lambda^n \pi_0 = \pi_n, \\ \lambda \bar{F}(n) \pi_n = \pi_{n+1}, \\ \lambda^2 \bar{F}(n) \bar{F}(n+1) \pi_n = \pi_{n+2}, \\ \vdots \\ \lambda^r \prod_{i=n}^{n+r-1} \bar{F}(i) \pi_n = \pi_{n+r} \end{cases} \quad (1)$$

根据平稳分布条件:

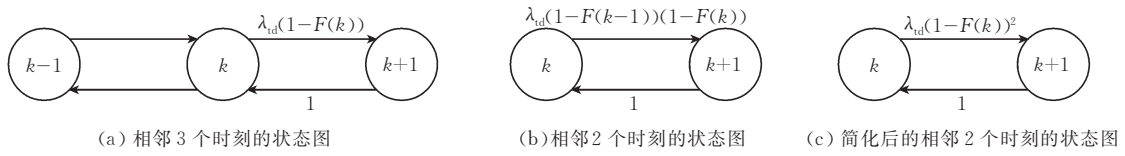


图 1 相邻时刻状态的近似性

图 2 是对 TD 和 RED 算法分别就类 TCP 的响应性流和类 UDP 的非响应性流进行数值模拟的结果,其中 RED 算法使用即时队列 (instantaneous queue) 和参数: $min=2, max=20, p_{max}=1$ 来决定包丢弃率,并且假设响应性流具有两种包到达速率:包丢失后的泊松到达平均速率 λ_{to} 和没有包丢失的泊松到达平均速率 λ_{td} ,并根据定理 2,有 $\lambda_{to}=0$. 在图 2(a)中,RED 在非响应性流下的丢包率和 TD 差不多,在轻负载下还高于 TD,即为保持较低队列长度的目的而牺牲了丢包率,但由于非响应性源没有流控,这种丢包策略实际上没有意义,并浪费了可用队列长度(见图 2(b))的资源.而对于响应性流效果则明显不同,RED 表现出很好的控制时延和减少丢包的性能.特别是在有较高系统负载 ($\lambda_{td}=1.5 \sim$

这样就得到 π_0 的值,代入式(1)进而推算出 π_i 各值.可以得到平均系统队列长度和平均包丢弃率:

$$\sum_{i=0}^{n+r} \pi_i = 1 \quad (2)$$

$$E[X] = \sum_{i=0}^{n+r} i \pi_i \quad (3)$$

$$E[F] = \sum_{i=n}^{n+r} F(i) \pi_i \quad (4)$$

对于响应性流的建模较为复杂,需要做两个假设性简化:

(1) 根据定理 2,设 $\lambda_{to}=0$,则拥塞避免阶段的泊松到达参数为 $\lambda_{td}(1-p)$,其中 $p=F(n)$;

(2) 相邻状态具有近似性. 非响应性流的状态图如图 1(a)所示,易得 $\lambda_{td}(1-F(k))\pi_k = \pi_{k+1}$, π_i 为达到平稳状态后的稳态概率. 定理 1 证明了响应性流过程具有二阶马尔可夫性,即 k 状态的流输入参数 λ_{td}^* 受到 $k-1$ 状态的影响,当 $k-1$ 状态执行过包丢弃时, k 状态的 $\lambda_{td}^* = \lambda_{to} = 0$,未执行过丢弃时,才会有 k 状态的 $\lambda_{td}^* = \lambda_{td}$,故 $\lambda_{td}^* = \lambda_{td}(1-F(k-1))$,如图 1(b). 考虑相邻两个状态队列长度的变化很小,当 λ_{td} 值较小而 $max \gg min$ 时,则 $F(k-1) \approx F(k)$,图 1(b)的状态可转化为图 1(c),这样就将模型简化为可变输入率的一阶马尔可夫链,按照前述方法可以得到平稳状态概率及相关参数. 对 TD(Tail Drop)算法可使用 $M/M/1/m$ 的排队模型.

2.5)时,两种流输入下的 TD 都长期处于队列高占有率的情况,其频繁的丢包导致了系统稳定性的急剧下降,但 RED 利用较小的丢包代价(相对 TD 的丢包率)使响应性源的流控机制成功作用于流输入率,有效地控制了队列长度(保证时延).但在较轻系统负载 ($\lambda_{td}=0.5 \sim 1$) 时,RED 的性能并不占优,这是由于 RED 对 min 以上的队列长度存在丢包敏感性,而 TD 只对达到 max 的队列长度存在丢包敏感性,在轻负载下,系统队列达到 max 的稳态概率很小,但达到 min 的概率则要大得多.总体上与 TD 相比,RED 在不增加甚至减少丢包率的前提下显著降低了系统时延.从图中还可以看出,对同一种算法,非响应性流和响应性流的效果也存在很大差别,说明响应性流对增强流的性能是有现实意义的.

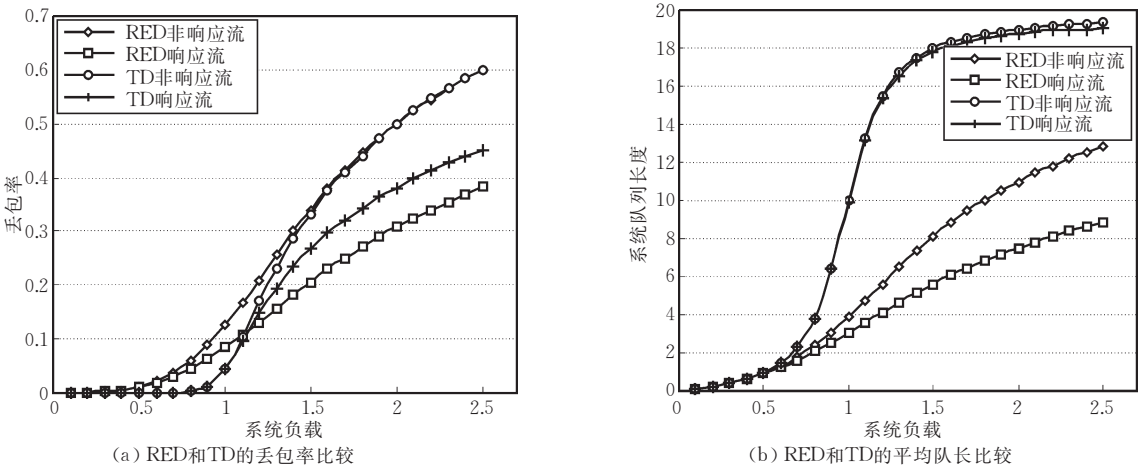


图 2 RED 和 TD 的性能比较

3 平均队列长度的分析

在上述模型中,RED 的丢弃率是建立在即时队列(instantaneous queue)的基础上,而实际上 RED 采用平均队列长度的概念并使用指数平滑(EWMA)的方法加以计算,但加入 EWMA 会使模型失去马尔可夫性,即使对其作矩形窗滑动的处理,也免不了高阶马尔可夫的复杂计算.文献[6]通过对突发性流量建立简化模型的基础上评估了平均队列长度的性能,指出平均队列长度确实避免了对突发源的丢包偏见并显著降低了突发时间段的丢包率,文中没有说明采用 EWMA 的理论依据以及其性能度量.我们认为 EWMA 在对突发源的处理上确实有一定效果,但它并不见得是一个针对具有 AIMD 行为流的好的平均队列长度计算方法.实际上,可以将平均队列长度看作是对下一个时刻队列长度的一步预测.

定理 3. EWMA 是时间序列模型 ARMA(1,1) 的一种特殊情况^[11].

ARMA(1,1)的形式是 $\tilde{X}_t = \phi X_t + a_t - \theta a_{t-1}$,取条件期望预测方法:

$$\begin{aligned}\tilde{X}_t &= \phi X_t - \theta a_{t-1} = \phi X_t - (1 - \lambda)(X_t - \tilde{X}_{t-1}) \\ &= \tilde{X}_{t-1} + \lambda(X_t - \tilde{X}_{t-1}) + \phi X_t - X_t.\end{aligned}$$

当 $\phi=1$ 时,就是 EWMA,即 EWMA 是 ARMA(1, 1)在 $\phi=1$ 时的一种特殊情况.

这就使得我们可以将 EWMA 纳入时间序列模型的框架中进行研究,用时间序列的方法和理论来评估 EWMA.观察 EWMA 的预测递推式 $\tilde{X}_t = X_t - \theta(X_t - \tilde{X}_{t-1})$,其预报结果是当前队列长度加上一个加权的误差修正,在一个具有二阶马尔可夫输入序列下考察该模型预报的合理性,它忽略了以

下几点:

(1)二阶相关性.定理 1 已经证明当前队列长度与前两个时刻的队列长度相关,因此预测递推式的 AR 部分应含有 X_{t-1} 和 X_t 因子,而 EWMA 只含后者;

(2)时变性.输入的随机性决定了参数的时变性,在 burst traffic 和 smooth traffic 交错的情况下,固定参数对远近历史的相关显然会顾此失彼;

(3)数据饱和.随着时间的增加,时变参数的变化却越来越不精确,原因在于过多的历史数据将新的数据淹没而反映不出时变特性,这就要求剔除过去无用的历史数据.

实际上,问题的实质就在于预报的准确性,合理的丢弃率算式应能跟踪并追随流的 AIMD 行为,这样才能在端到中继再到端的链路产生流量协调一致的波动.基于上述分析,改写平均队列长度计算方法,将 X_{t-1} 因子加入预测递推式的 AR 部分,用 ARMA(2,1)算式代替 EWMA.

4 带 EWMA 的 RED 算法及 ARMA 算法的通用模型

由于 EWMA 的算子包含过去任何时刻的状态,使前述模型因失去了无后效性而不能成立.然而,如果权值 λ 非常小,那么各个历史因子的权值变化就非常小,而且递推值的变化非常缓慢,而 EWMA 的权值之和等于 1,这样递推值近似于所有历史因子的均值,也就是稳态下队列长度的期望值.尽管只要历史足够长,EWMA 权值最终成曲线分布,但对历史进行分段处理,则在一定段内该假设近似是成立的.这样,在一定输入下,系统丢弃率是一个固定值,设队列期望值为 $E(X)$,即得

$$F(n) = p_{\max} \times (E(X) - \min) / (\max - \min),$$

$$\min < n < \max \quad (5)$$

尽管恢复了马尔可夫性,但 $E(X)$ 是个未知量,根据式(2),通过联立方程可以求此值:

$$\begin{cases} \lambda \pi_0 = \pi_1, \lambda^2 \pi_0 = \pi_2, \dots, \lambda^n \pi_0 = \pi_n, \\ \lambda \bar{F}^2(n) \pi_n = \pi_{n+1}, \\ \lambda^2 \bar{F}^2(n) \bar{F}^2(n+1) \pi_n = \pi_{n+2}, \\ \vdots \\ \lambda^r \prod_{i=n}^{n+r-1} \bar{F}^2(i) \pi_n = \pi_{n+r} \sum_{i=n}^{n+r} i \pi_i = E(X) \end{cases} \quad (6)$$

求得 $E(X)$ 之后,相应的目标参量随之被解出.在实际应用中权值 λ 的设置是较小的,如文献[12]推荐为 0.002.

接下来考虑 ARMA(2,1)的预报计算模型,为便于建模和对比效果,我们采用和 EWMA 近似的参数设置,即设权值 λ 非常小,同时根据 EWMA 递推式 $\tilde{X}_t = X_t - \theta(X_t - \tilde{X}_{t-1})$,设置 ARMA(2,1)的递推式 $\tilde{X}_t = \phi_1 X_t + \phi_2 X_{t-1} - \theta(X_t - \tilde{X}_{t-1})$ 的 AR 参数 $\phi_1 = \phi_2 = 1/2$,它比 EWMA 在 AR 部分多了上上时刻状态的因子,这使模型又失去了马尔可夫性.

定理 4. 在上述条件下,ARMA(2,1)在 t 时刻的预报可近似表达为 $3E(X)/2 - X(t)/2$.

推理如下:

$$\begin{aligned} \tilde{X}_t &= X_t/2 + X_{t-1}/2 - (1-\lambda)(X_t - \tilde{X}_{t-1}) \\ &= (X_{t-1} - X_t)/2 + \lambda X_t + (1-\lambda)\tilde{X}_{t-1} \\ &= (X_{t-1} - X_t)/2 + \lambda X_t + \\ &\quad (1-\lambda)(X_{t-2} - X_{t-1})/2 + \\ &\quad \lambda(1-\lambda)X_{t-1} + (1-\lambda)^2\tilde{X}_{t-2} \\ &= (X_{t-1} - X_t)/2 + \\ &\quad (1-\lambda)(X_{t-2} - X_{t-1})/2 + \dots + \\ &\quad (1-\lambda)^n(X_{t-n-1} - X_{t-n})/2 + \dots + \\ &\quad \sum_{i=0}^n \lambda(1-\lambda)^i X_{t-i} \\ &= -X_t/2 + \lambda X_{t-1}/2 + \lambda(1-\lambda)X_{t-2}/2 + \dots + \\ &\quad \lambda(1-\lambda)^n X_{t-n-1}/2 + \dots + E(X) \\ &= -X_t/2 + E(X)/2 + E(X) \\ &= 3E(X)/2 - X_t/2. \end{aligned}$$

所求得证.

同 RED 模型一样,马尔可夫性得以恢复,同时得到各状态系统丢弃率,设队列期望值为 $E(X)$,即得

$$F(n) = p_{\max} \times (3E(X)/2 - n/2 - \min) /$$

$$(\max - \min), \quad \min < n < \max \quad (7)$$

但 $E(X)$ 也是个未知量,同理根据式(6)求得 $E(X)$,相应的目标参量随之被解出.若使 $\lambda=1$,则

ARMA(2,1)转化为 AR(2)模型,该模型未被考察的原因是其对应于 ARMA(2,1),就如同即时队列 RED 对应于 EWMA-RED,其区别仅是对 burst traffic 的平滑效果不同.实际上在得到上述三者的性能曲线后,AR(2)的性能曲线不难被推测出来.

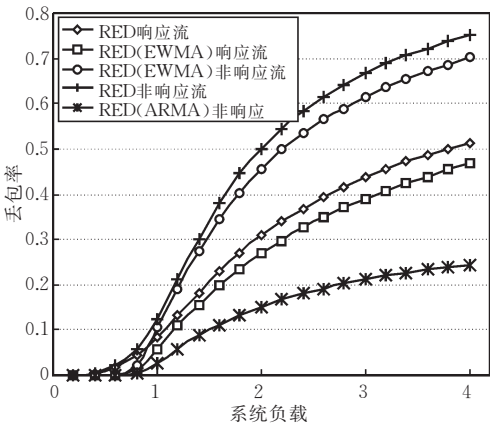
5 算法性能曲线的比较

采用等同图 2 的参数,分别对 Instantaneous queue-RED(以下简称 RED),EWMA-RED 和 ARMA-RED 进行数值仿真,结果如图 3 所示.我们发现无论是对响应性流还是非响应性流,EWMA-RED 在总体性能上略高于 RED,在输入参数 λ 超过一定值(对响应性流约为 3,对非响应性流约为 2)后,EWMA-RED 在丢包率和系统队列长度上都略低于 RED,但在此之前,EWMA-RED 的平均队长却略高于 RED.分析原因是当系统负载较轻,或突发源所占比重较小时,由于 EWMA 带有历史数据,其对突发源和非突发源的公平对待及对突发流一定的容纳能力,故允许突发流在一定程度上存在,而相比之下,RED 对突发流的强烈抵制几乎完全抑制了突发源的流量,结果使得 EWMA-RED 的平均系统队列更高一些.值得一提的是,在轻负载状态下刻意维持低平均系统队列并不见得是好事,这降低了系统资源的利用率.当系统负载较重,或突发源所占比重较大时,大多数时刻状态下的即时队列值都高于 EWMA 的历史均值,于是 RED 的丢包率明显高于 EWMA-RED.因此负载越重则 EWMA 的优越性能越能体现.从上述分析可以推断,理想的性能曲线应该具有的特征是:(1)平均丢包率越小越好,这是系统稳定性的一个重要标志;(2)系统队列长度在低负载下应该高一些,而在重负载下尽可能低一些,或者说低负载下的系统队列长度尽可能地接近于重负载下的队列长度.

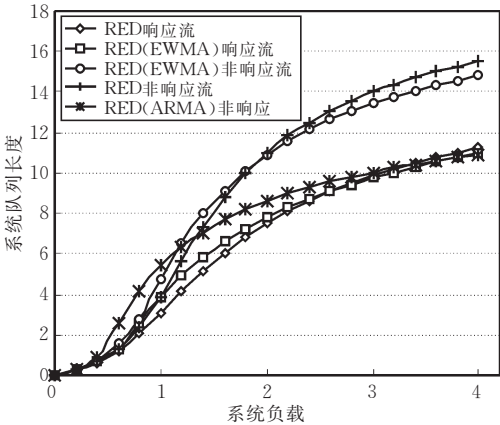
我们考察 ARMA-RED 的表现,显然其曲线更加接近于理想的性能曲线.与 EWMA-RED 对比 RED 不同,从一开始 ARMA-RED 的丢包率就比其他算法低了一个数量级,并且在低负载下系统资源的利用也很充分.随着负载的增大,其队列长度的曲线开始逼近 EWMA-RED 和 RED,并在约 3.5 处相交,随后 ARMA 曲线值成为所有算法曲线的最低值,而此时的丢包率只有 EWMA-RED 的一半.这从侧面印证了我们的推测,ARMA 应该在任何负载下的长期预报都比 EWMA 准确,相比之下,EWMA-RED 之所以在轻负载下的丢包率偏高,潜在原因

就在于 EWMA 在轻负载下的预报准确度还不如即时队列,因为此时时变输入量之间的变化量很小,即时队列便可看作一种简单的预报方法且效果尚可.

不过在时延抖动上,RED 要优于 EWMA-RED 和 ARMA-RED,原因是 RED 是即时的,而后两者都带有滞后性.



(a) EWMA和ARMA的丢包率比较



(b) EWMA和ARMA的平均队长比较

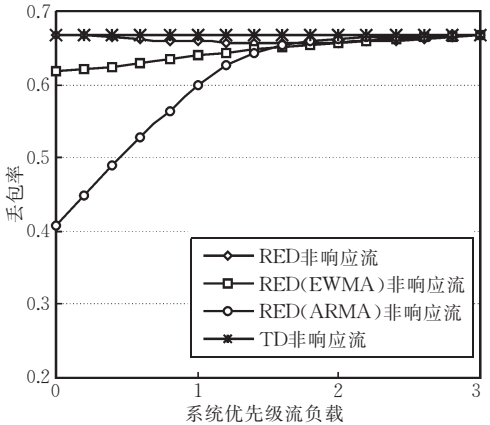
图 3 EWMA 和 ARMA 的性能比较

6 支持优先级的丢弃策略

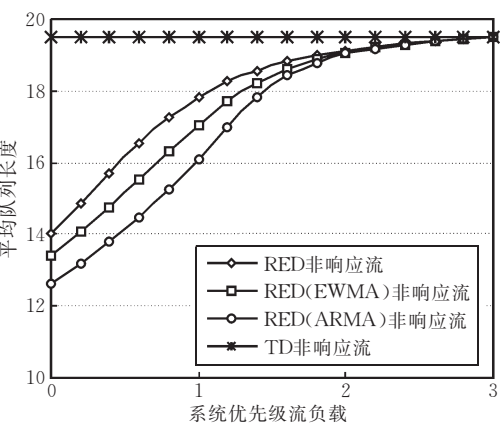
前述的算法从量的角度去解决每个流的丢包问题,但对于需要优先级支持的网络流(如分层组播网络)缺乏了质方面的考虑,将优先级丢弃机制与普通单播网络的队列管理策略结合起来是一个有效的解决方案.这种方案以平时的丢包来保证有足够的空间容纳高优先级的突发性包,达到保护高优先级包的目的.模型采用两级优先级方案,对低优先级包使用通用的单播队列管理策略,对高优先级包使用TD丢弃策略,这样高优先级包被尽可能地接收直至缓冲溢出,模型不考虑对缓冲进行排队及将高优先级包替换低优先级包的情况.设为低优先级包服从参数 λ_1 泊松到达,高优先级包服从参数 λ_2 泊松到达,改写式(1)可得到

$$\left\{ \begin{array}{l} \hat{\lambda}\pi_0=\pi_1, \hat{\lambda}^2\pi_0=\pi_2, \cdots, \hat{\lambda}^n\pi_0=\pi_n, \hat{\lambda}=\lambda_1+\lambda_2 \\ \check{\lambda}\bar{F}(n)\pi_n=\pi_{n+1}, \\ \check{\lambda}^2\bar{F}(n)\bar{F}(n+1)\pi_n=\pi_{n+2}, \\ \vdots \\ \check{\lambda}^r\prod_{i=n}^{n+r-2}\bar{F}(i)\pi_n=\pi_{n+r-1}, \\ (\check{\lambda}^r\prod_{i=n}^{n+r-1}\bar{F}(i)+\lambda_2/\hat{\lambda})\pi_n=\pi_{n+r}, \check{\lambda}=\lambda_1/\hat{\lambda} \end{array} \right. \quad (8)$$

使用图 2 一样的参数,另设 $\lambda_1+\lambda_2=3$,得到如图 4 的数值模拟结果,横坐标是高优先级流的泊松到达参数值.可以看到,随着高优先级流比重的增加,各算法性能随之变坏,当两种优先级流各占一半时,各种算法性能开始迅速趋向一致,当高优先级流比重少于一半时,各算法的性能有较明显的差距.在对平均队列长度的控制上,TD 以外的三种策略都有较好的表现,从丢包率上看,RED 为承担高优先



(a) 各算法的丢包率比较



(b) 各算法的平均队长比较

图 4 优先级队列中各算法的性能比较

级包的容纳能力,对低优先级包采用了接近 TD 的丢弃率,而 EWMA 和 ARMA 则有相对富裕的回旋余地.从图中也可以看出,EWMA 和 ARMA 的两个性能曲线随着高优先级流比重的增加量都几乎是线性的,因此不是好的曲线走势,这就提出了新的研究课题,在面对不同应用服务的要求(如有的要求高可靠,有的要求低时延)的情况下,如何改良算法以使曲线走势更能服从不同应用的需求.

7 结束语

通过本文研究可知,RED 算法的稳定性和控制低平均时延能力要明显高于 TD 算法,而 EWMA-RED 的性能主要体现在对突发的平滑能力及对突发源的公平性,但在可靠性和平均时延上并未比 RED 有太好的表现,潜在原因就是采用 EWMA 计算平均队列长度对提升这两种性能并无太大帮助,ARMA-RED 则对系统输入具有较好的预测能力,因而明显地提升了稳定性和降低了平均时延,且不影响对突发的平滑.本文的研究也表明了从时序推断的角度来分析研究拥塞控制策略的价值.

参 考 文 献

- [1] Chung J, Claypool M. Analysis of active queue management//Proceedings of the 2nd IEEE International Symposium on Network Computing and Applications (NCA'03). Cambridge, MA, USA: IEEE CS Press, 2003: 359-366
- [2] Braden B, Clark D, Crowcroft J, Davie B, Deering S. Recommendation on queue management and congestion avoidance in the internet. Internet Engineering Task Force, Request

for Comments (RFC)2309, 1998

- [3] Hasegawa G, Kurata K, Murata M. A study on high speed packet scheduling method for achieving fairness among TCP connections. *Journal of High Speed Networks*, 2003, 12: 29-49
- [4] Mokhtar A, Azizoglu M. A random early discard framework for congestion control in ATM networks//Proceedings of the IEEE ATM Workshop. Kochi, Japan, 1999: 45-50
- [5] Alazemi H, Mokhtar A, Azizoglu M. Stochastic approach for modeling of random early detection gateways in TCP/IP networks//Proceedings of the IEEE International Conference on Communications (ICC). Helsinki, Finland, 2001: 1747-1751
- [6] Bonald T, May M, Bolot J C. Analytic evaluation of RED performance//Proceedings of the IEEE INFOCOM 2000. Tel-Aviv, Israel, 2000: 1415-1424
- [7] Misra V, Gong W B, Towsley D. Fluid-based analysis of a network of AQM routers supporting TCP flows with an application to RED//Proceedings of the ACM SIGCOMM 2000. Stockholm, Sweden, 2000: 151-160
- [8] Kamal A E. Modeling TCP Reno with RED-based routers. *ATM SIGMETRICS Performance Evaluation Review*, 2002, 30(3): 23-25
- [9] Zhang Z S, Hasegawa G, Murata M. Performance analysis and improvement of HighSpeed TCP with TailDrop/RED routers//Proceedings of the IEEE Computer Society's 12th Annual International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunications Systems (MASCOTS'04). Netherlands, 2004: 505-512
- [10] Padhye J, Firoiu V, Towsley D, Kurose J. Modeling TCP throughput: A simple model and its empirical validation//Proceedings of the ACM SIGCOMM'98. Vancouver, Canada, 1998: 303-314
- [11] Pandit S M, Wu S M. Time Series and System Analysis with Applications. John Wiley & Sons, 1983
- [12] Floyd S, Jacobson V. Random early detection gateways for congestion avoidance. *IEEE/ACM Transactions on Networking*, 1993, 1(4): 397-413



XU Li-Bo, born in 1976, Ph.D. candidate. His research interests include QoS and P2P network.

WU Guo-Xin, born in 1956, professor. His research interests include QoS and P2P network.

Background

To avoiding congestion collapse on Internet links, relay devices such as routers adopt AQM to tolerate bursting flows. IETF introduces RED to design AQM arithmetic. To combing end-user activity to RED, there appear several new AQM schemes to improve RED control ability. For no modeling expression, majority of these schemes lack theoretic explication and deduction.

The main contribution of this paper is modeling and ana-

lyzing various AQM strategies. Through Markov chain description, the model of RED behaves under unresponsive and responsive end-flows is built. Then the nature and expression of average queue length in RED are observed. At last a new arithmetic with time series ARMA for forecasting average queue length is given and the simulation results show its obvious advantages on reliability and resource utilization.