

# 一种自相似业务量预测的卡尔曼滤波算法

郭庆, 那振宇, 顾学迈

(哈尔滨工业大学通信技术研究所, 150001, 哈尔滨)

**摘要:** 针对网络拥塞控制中不能准确预测自相似业务量的问题, 提出了一种噪声在线估计卡尔曼滤波(NOEFK)算法. NOEFK 算法不依赖于业务源反馈信息, 通过观测节点处当前和过去时刻的业务量来预测下一时刻的业务量, 并建立了业务量的卡尔曼滤波状态方程和观测方程, 给出了递推形式的状态向量最佳估计形式. 考虑到未知状态方程和观测方程噪声的统计特性, 采用在线估值法, 并引入遗忘因子对噪声的统计特性进行估计. NOEFK 算法预测准确、偏差小. 仿真结果表明, 与经典卡尔曼滤波算法和时间序列预测方法比较, NOEFK 算法能够更精确地预测自相似业务量, 预测误差可降低 60% 以上.

**关键词:** 卡尔曼滤波; 自相似业务; 预测算法

**中图分类号:** TN915.01 **文献标志码:** A **文章编号:** 0253-987X(2009)04-0057-05

## A Kalman Filtering Algorithm for Self-Similar Traffic Prediction

GUO Qing, NA Zhenyu, GU Xuemai

(Communication Research Center, Harbin Institute of Technology, Harbin 150001, China)

**Abstract:** A noise on-line estimation Kalman filtering (NOEFK) algorithm is presented to deal with the inaccurate self-similar traffic prediction in network congestion control. The proposed algorithm is independent of the feedback information from traffic sources, and predicts the traffic through observing both the current and previous traffics in a node. Both the state equation and the observation equation are established, and then an optimal recursive formula for the estimation of the state vector is given. By taking the unknown noise statistics of both the state equation and the observation equation into account, an on-line estimation method with forgetting factor is used to estimate the noise statistics. Comparisons with existing algorithms show that the NOEFK algorithm has the advantages of high accuracy and minor prediction error. Simulation results show that the NOEFK algorithm predicts self-similar traffic more accurately than the classical Kalman filtering and time series prediction algorithms do, and that the prediction error is reduced by more than 60%.

**Keywords:** Kalman filtering; self-similar traffic; prediction algorithm

近年来对局域网和广域网的分析表明, 网络流量普遍存在自相似性<sup>[1-2]</sup>. 自相似性会使得网络流量在很大的时间尺度范围内突然增加或减少. 这种流量突发性会使得网络节点处队列性能严重恶化, 一方面降低了资源利用率, 另一方面使得时延和抖动增加, 无法满足实时性业务的服务质量. 因此, 对网

络节点处到达的业务量进行实时准确预测, 是改善网络服务质量的有效方法.

目前常用的预测方法有神经网络预测<sup>[3-5]</sup>、小波预测<sup>[6-7]</sup>和时间序列预测<sup>[8-10]</sup>. 神经网络预测不依赖先验知识, 具有自适应学习的能力, 但却需要进行大量迭代运算来不断修正模型, 增加了时间和空间的

复杂性. 小波预测结果较好, 但是实时性较差. 以 FARIMA 模型为代表的时间序列方法能同时描述业务的短相关和长相关性, 但是参数确定依赖于对流量特性的先验知识<sup>[9]</sup>, 实现较复杂. 文献[11-12]提出使用卡尔曼滤波器预测业务量. 文献[11]基于泊松过程模型提出卡尔曼滤波方程, 这不能很好地适应自相似业务流, 并且状态转移矩阵和测量矩阵的确定依赖于业务源反馈信息来实现, 在大时延网络中, 这种反馈会严重影响实时性. 文献[12]在文献[11]模型的基础上应用小波进行预测, 但预测精度没有明显提高.

本文基于观测到达业务量建立卡尔曼滤波算法的观测方程和状态方程, 由于状态方程和观测方程各自噪声的统计特性未知, 因此采用在线估计方法对噪声进行估计, 引入遗忘因子, 使估值器逐渐减小陈旧数据的权重, 得到了很好的预测结果.

## 1 卡尔曼滤波预测模型

经典卡尔曼滤波方程由描述状态向量的状态方程和描述观测向量的观测方程共同表示, 即

$$\mathbf{X}(n+1) = \mathbf{F}(n+1, \mathbf{X}_n)\mathbf{X}(n) + \mathbf{w}(n) \quad (1)$$

$$\mathbf{y}(n) = \mathbf{C}(n)\mathbf{X}(n) + \mathbf{v}(n) \quad (2)$$

式中:  $\mathbf{X}(n)$  是  $n$  时刻的  $M \times 1$  维状态向量;  $\mathbf{F}(n+1, n)$  是  $M \times M$  维状态转移矩阵;  $\mathbf{w}(n)$  是  $M \times 1$  维状态噪声向量;  $\mathbf{y}(n)$  是  $N \times 1$  维观测向量;  $\mathbf{C}(n)$  是  $N \times M$  维观测矩阵;  $\mathbf{v}(n)$  是  $N \times 1$  维观测噪声向量.

考虑到在节点处到达业务量是可观测的, 设  $T$  为观测时间间隔. 令  $x(n)$  表示  $n$  时刻到达的业务量, 那么  $x'(n)$  表示  $x(n)$  的一阶微分. 如果时间间隔  $T$  不太大, 那么可以用一阶差分方程得到如下关系

$$x(n) = x(n-1) + Tx'(n-1) \quad (3)$$

式中:  $x'(n)$  是流量变化率, 即流速. 同理, 可得到

$$x''(n) = x'(n-1) + Tx''(n-1) \quad (4)$$

式中:  $x''(n)$  是流速变化率. 很多原因可以导致节点处业务流量的速率发生变化, 如信道状态的随机性、用户访问网络的不确定性、业务持续时间的重尾分布特性以及各种干扰等. 考虑将这些因素视为导致流速变化的统计特性未知的噪声  $w(n)$ , 因而

$$x''(n-1) = w(n) \quad (5)$$

令状态向量

$$\mathbf{X}(n) = \begin{bmatrix} x(n) \\ x'(n) \\ x''(n) \end{bmatrix} \quad (6)$$

则上面的式(3)~式(5)可以写作

$$\mathbf{X}(n) = \mathbf{F}(n, n-1)\mathbf{X}(n-1) + \mathbf{w}(n) \quad (7)$$

式中:  $\mathbf{F}(n, n-1)$  是状态转移向量(简记为  $\mathbf{F}$ ), 即

$$\mathbf{F} = \begin{bmatrix} 1 & T & 0 \\ 0 & 1 & T \\ 0 & 0 & 0 \end{bmatrix} \quad (8)$$

状态噪声向量为

$$\mathbf{w}(n) = \begin{bmatrix} 0 \\ 0 \\ \tau w(n) \end{bmatrix} \quad (9)$$

此外, 观测方程为

$$\mathbf{y}(n) = \mathbf{C}(n)\mathbf{X}(n) + \mathbf{v}(n) \quad (10)$$

式中:  $\mathbf{C}(n) = [1 \ 0 \ 0]$  表示观测矩阵;  $\mathbf{v}(n)$  表示观测噪声.

下面分析所提滤波问题的解法. 由式(8)知, 状态转移矩阵  $\mathbf{F}$  不可逆, 因而无法用经典卡尔曼滤波法进行递推计算. 考虑根据式(7)和式(10), 在未知  $\mathbf{w}(n)$  和  $\mathbf{v}(n)$  的情况下, 对状态向量  $\mathbf{X}(n)$  有 2 种估计式. 根据式(7), 可得到  $\mathbf{X}(n)$  的第 1 种估计式, 即

$$\hat{\mathbf{X}}_1(n) = \mathbf{F}\hat{\mathbf{X}}(n-1) \quad (11)$$

由式(10)可得到  $\mathbf{X}(n)$  的第 2 种估计式, 即

$$\mathbf{y}(n) = \mathbf{C}\hat{\mathbf{X}}_2(n) \quad (12)$$

考虑采用 2 种估计式的最佳线性组合来确定  $\mathbf{X}(n)$ . 设  $\mathbf{A}(n)$  是最佳加权矩阵, 则

$$\hat{\mathbf{X}}(n) = [\mathbf{I} - \mathbf{A}(n)]\hat{\mathbf{X}}_1(n) + \mathbf{A}(n)\hat{\mathbf{X}}_2(n) \quad (13)$$

式中:  $\mathbf{A}(n)$  为  $M \times M$  维.

为求  $\mathbf{A}(n)$ , 令

$$\mathbf{A}(n) = \mathbf{m}(n)\mathbf{C}(n) \quad (14)$$

这样, 就将求解最佳系数矩阵  $\mathbf{A}(n)$  的问题变为求最佳矩阵  $\mathbf{m}(n)$  的问题.

将式(14)代入式(13)并整理, 则状态向量的估计误差向量  $\xi(n)$  可以表示为

$$\begin{aligned} \xi(n) &= \mathbf{X}(n) - \hat{\mathbf{X}}(n) = [\mathbf{I} - \mathbf{m}(n)\mathbf{C}(n)][\mathbf{X}(n) - \hat{\mathbf{X}}_1(n)] + \mathbf{m}(n)[\mathbf{C}(n)\mathbf{X}(n) - \mathbf{y}(n)] = \\ &= [\mathbf{I} - \mathbf{m}(n)\mathbf{C}(n)][\mathbf{X}(n) - \hat{\mathbf{X}}_1(n)] - \mathbf{m}(n)\mathbf{v}(n) \end{aligned} \quad (15)$$

确定  $\mathbf{A}(n)$  的原则是, 使得  $\mathbf{X}(n) - \hat{\mathbf{X}}(n)$  的均方值最小. 考察误差向量  $\xi(n)$  的相关矩阵  $\mathbf{R}(n)$ , 并考虑到噪声  $\mathbf{v}(n)$  与  $\mathbf{X}(n) - \hat{\mathbf{X}}_1(n)$  彼此不相关, 得

$$\begin{aligned} \mathbf{R}(n) &= [\mathbf{I} - \mathbf{m}(n)\mathbf{C}(n)]E[(\mathbf{X}(n) - \hat{\mathbf{X}}_1(n)) \cdot \\ & \quad (\mathbf{X}(n) - \hat{\mathbf{X}}_1(n))^T][\mathbf{I} - \mathbf{m}(n)\mathbf{C}(n)]^T + \\ & \quad \mathbf{m}(n)\sigma^2 \mathbf{m}^T(n) \end{aligned} \quad (16)$$

式中: $\sigma^2$  为现测过程方差. 令

$$\mathbf{R}_1(n) = E[(\mathbf{X}(n) - \hat{\mathbf{X}}_1(n))(\mathbf{X}(n) - \hat{\mathbf{X}}_1(n))^T] \quad (17)$$

则式(16)可以记为

$$\mathbf{R}(n) = [\mathbf{I} - \mathbf{m}(n)\mathbf{C}(n)]\mathbf{R}_1(n) \cdot [\mathbf{I} - \mathbf{m}(n)\mathbf{C}(n)]^T + \mathbf{m}(n)\sigma^2 \mathbf{m}^T(n) \quad (18)$$

令  $\frac{\partial \text{tr} \mathbf{R}(n)}{\partial \mathbf{m}(n)} = 0$ , 可以解得

$$\mathbf{m}(n) = \mathbf{R}_1(n)\mathbf{C}^T(n)[\mathbf{C}(n)\mathbf{R}_1(n)\mathbf{C}^T(n) + \sigma^2]^{-1} \quad (19)$$

将式(19)带入式(18), 整理得到

$$\mathbf{R}(n) = [\mathbf{I} - \mathbf{m}(n)\mathbf{C}(n)]\mathbf{R}_1(n) \quad (20)$$

在式(20)中, 需要计算  $\mathbf{R}_1(n)$ . 注意到

$$\mathbf{X}(n) - \hat{\mathbf{X}}_1(n) = \mathbf{F}(\mathbf{X}(n-1) - \hat{\mathbf{X}}_1(n-1)) + \mathbf{w}(n) \quad (21)$$

于是

$$\mathbf{R}_1(n) = E\{[\mathbf{F}(\mathbf{X}(n-1) - \hat{\mathbf{X}}_1(n-1)) + \mathbf{w}(n)] \cdot [\mathbf{F}(\mathbf{X}(n-1) - \hat{\mathbf{X}}_1(n-1)) + \mathbf{w}(n)]^T\} \quad (22)$$

利用状态转移矩阵  $\mathbf{F}$  与取期望无关, 同时  $\mathbf{w}(n)$

和  $\mathbf{x}(n)$  不相关, 则式(22)可以整理为

$$\mathbf{R}_1(n) = \mathbf{F}\mathbf{R}_1(n-1)\mathbf{F}^T + \mathbf{Q}_w \quad (23)$$

式中: $\mathbf{Q}_w$  是状态噪声的相关矩阵. 可见,  $\mathbf{R}_1(n)$  可以递推求得.

## 2 在线噪声估计的改进卡尔曼滤波

由于噪声的统计特性无法预先知道, 这里考虑利用观测数据, 使用在线估计方法自适应地估计噪声的统计特性. 由上面的分析可以看出, 利用上面卡尔曼滤波模型的时候, 需要用到状态噪声的相关矩阵  $\mathbf{Q}_w$  和测量过程的方差  $\sigma^2$ . Sage 和 Husa<sup>[13]</sup> 给出了在线估计噪声的方法. 令  $\mathbf{Q}_w$  代表状态噪声的相关矩阵,  $\mathbf{R}_v$  代表测量噪声的相关矩阵, 则它们的估值  $\hat{\mathbf{Q}}_w$  和  $\hat{\mathbf{R}}_v$  可以表示为

$$\hat{\mathbf{Q}}_w(n+1) = (1 - d_n)\hat{\mathbf{Q}}_w(n) + d_n \frac{1}{n+1} [\mathbf{K}(n+1)\boldsymbol{\varepsilon}(n+1)\boldsymbol{\varepsilon}^T(n+1)\mathbf{K}^T(n+1) + \mathbf{P}(n+1) - \mathbf{F}\mathbf{P}(n)\mathbf{F}^T] \quad (24)$$

$$\hat{\mathbf{R}}_v(n+1) = (1 - d_n)\hat{\mathbf{R}}_v(n) + d_n [\boldsymbol{\varepsilon}(n+1)\boldsymbol{\varepsilon}^T(n+1) - \mathbf{C}\mathbf{P}(n+1 | n)\mathbf{C}^T] \quad (25)$$

式中: $\mathbf{K}(n)$  是卡尔曼增益;  $\boldsymbol{\varepsilon}(n+1)$  是观测值与预测值的偏差同噪声均值的差量(即新息);  $\mathbf{P}(n+1 | n)$

是预测状态误差相关矩阵;  $\mathbf{P}(n)$  是滤波状态误差相关矩阵. 在式(24)和式(25)中, 考虑到了噪声的时变性, 因此采用加权的估计方法, 以加强新观测数据的权重, 其中  $d_k = (1-b)/(1-b^{k+1})$ ,  $b$  称为遗忘因子,  $0 < b < 1$ , 一般取  $b > 0.9$ .

本文中  $\hat{\mathbf{R}}_v$  矩阵退化为 1 行 1 列的  $\hat{\sigma}^2$ ,  $\boldsymbol{\varepsilon}(n+1)$  也退化为 1 行 1 列, 即  $\varepsilon(n+1)$ . 因而, 式(24)和式(25)可以化为以下两式

$$\begin{aligned} \hat{\mathbf{Q}}_w(n+1) &= (1 - d_n)\hat{\mathbf{Q}}_w(n) + \\ d_n \frac{1}{n+1} &[\varepsilon^2(n+1)\mathbf{K}(n+1)\mathbf{K}^T(n+1) + \\ &\mathbf{P}(n+1) - \mathbf{F}\mathbf{P}(n)\mathbf{F}^T] \end{aligned} \quad (26)$$

$$\begin{aligned} \hat{\sigma}^2(n+1) &= (1 - d_n)\hat{\sigma}^2(n) + \\ d_n &[\varepsilon^2(n+1) - \mathbf{C}\mathbf{P}(n+1 | n)\mathbf{C}^T] \end{aligned} \quad (27)$$

综上所述, 本文提出的改进卡尔曼滤波业务流量预测算法, 对于  $n=1, 2, \dots$ , 需要计算式(11)、式(19)、式(20)、式(23)、式(24)和式(27), 其中

$$\begin{aligned} \mathbf{K}(n+1) &= \mathbf{P}(n+1 | n)\mathbf{C}^T(n) \cdot \\ &(\mathbf{C}(n)\mathbf{P}(n+1 | n)\mathbf{C}^T(n) - \hat{\sigma}^2(n))^{-1} \end{aligned} \quad (28)$$

$$\varepsilon(n+1) = \mathbf{y}(n+1) - \mathbf{C}(n)\mathbf{F}\hat{\mathbf{X}}(n) - \hat{r}(n) \quad (29)$$

$$\mathbf{P}(n+1) = [\mathbf{I} - \mathbf{K}(n+1)\mathbf{C}(n)]\mathbf{P}(n+1 | n) \quad (30)$$

$$\mathbf{P}(n+1 | n) = \mathbf{F}\mathbf{P}(n)\mathbf{F}^T + \hat{\mathbf{Q}}_w(n) \quad (31)$$

$$\begin{aligned} \hat{r}(n+1) &= (1 - d_n)\hat{r}(n) + \\ d_n &[\mathbf{y}(n+1) - \mathbf{C}(n)\hat{\mathbf{X}}(n+1)] \end{aligned} \quad (32)$$

式中: $\hat{r}(n)$  为观测噪声均值, 即  $E[\mathbf{v}(n)]$ .

## 3 实验结果

以 Bellcore 研究中心收集的实测数据为依据, 对本文提出的噪声在线估计卡尔曼滤波(NOEFK)预测算法进行了仿真, 并且与经典卡尔曼滤波(CKF)算法和文献[9]中的 FARIMA 算法进行了比较, 其中 FARIMA 模型参数的选取见文献[9]. 在 CKF 算法中, 假定状态噪声和测量噪声的方差是 1, 取遗忘因子  $b=0.95$ , 仿真中预测时长为 5 s. 为了便于对照, 图 1 和图 2 中还给出了以 0.05 s 为采样间隔的真实业务量的曲线.

从图 1 可以看出, 本文算法对真实业务量的预

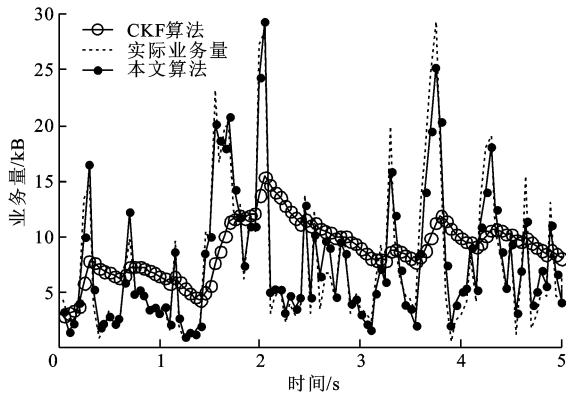


图 1 本文算法与 CKF 算法业务量预测性能

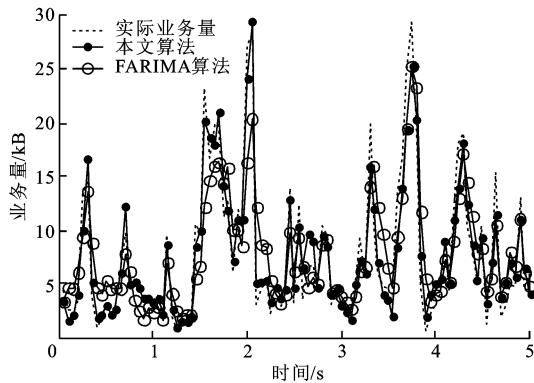


图 2 本文算法与 FARIMA 算法业务量预测性能

测效果很好,而 CKF 算法预测效果很差. 尽管 CKF 算法预测结果很不准确,但是对照真实流量曲线可以发现,CKF 算法的预测趋势同真实流量的变化几乎是一致的,但它的估计过于保守,原因是 CFK 算法没有准确估计噪声,而本文算法能够对噪声做出准确估计,因而对流量的预测更准确.

从图 2 可以看出,本文算法的业务量预测性能要优于 FARIMA 算法. 这是因为利用 FARIMA 模型预测的前提是要求时间序列是广义平稳过程,但是实际的网络流量往往无法符合这样的条件. 当网络流量发生剧烈变化的时候,FARIMA 模型往往不能跟踪流量的快速变化(这从图 2 中可以看出),而本文算法由于对噪声的统计特性做出了估计,因而能够较为及时、准确地反映出网络流量变化的趋势,流量预测更准确.

从图 3 可以看出,本文算法的预测相对误差较小. 经计算,本文算法的预测误差比 FARIMA 算法降低了 60% 以上,曲线较少出现大的峰值. 这说明本文算法预测的业务量跟真实值之间更接近,预测精度更高. 这里的相对误差是指预测值同真实值之

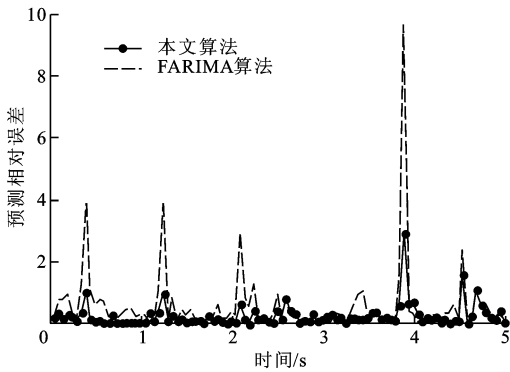


图 3 本文算法与 FARIMA 算法业务量预测误差比较

差的绝对值与真实值的比值,即

$$E_0 = \frac{|T_p - T_r|}{T_r} \tag{33}$$

式中:  $T_p$  为算法预测业务量;  $T_r$  为实际业务量.

表 1 还给出了本文算法和 FARIMA 算法的运算时间,可以看出本文算法所用时间不到 CKF 算法的 50%,因而计算量较小,实时性更高.

表 1 本文算法和 FARIMA 算法的运算时间比较

| 参数     | 本文算法 | FARIMA 算法 |
|--------|------|-----------|
| 运算时间/s | 0.82 | 1.71      |

4 结 论

本文提出了一种自相似业务量预测的卡尔曼滤波算法——NOEKF 算法. 该算法利用观测的业务量,推导卡尔曼滤波的状态方程和观测方程,并给出了算法的递推过程,同时对噪声进行实时估计,利用遗忘因子,减少陈旧数据的权重,以提高业务量急剧变化的部分的预测精度. 实验结果表明,本文算法具有较高的预测精度. 由于算法计算步骤是递推的,只需要存储上一时刻的值,因而计算量较小. 同时,算法仅根据当前和过去时刻的业务量进行预测,不需要源和宿之间的反馈,从而有利于减小时延. 本文的 NOEKF 算法适用于大时延网络环境下对业务流量的预测,为解决拥塞控制和优化网络管理提供了保障.

参考文献:

[1] WILLINGER W, TAQQU M S, SHERMAN R, et al. Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level [J]. IEEE/ACM Trans on Networking, 1997, 5(1): 71-86.

[2] CROVELLA M E, BESTAVROS A. Self-similarity in world wide web traffic evidence and possible causes [J]. IEEE/ACM Trans on Networking, 1997, 5(6):

- 835-846.
- [3] 王兆霞, 孙雨耕, 陈增强, 等. 基于模糊神经网络的网络业务量预测研究 [J]. 通信学报, 2005, 26(3): 136-140.  
WANG Zhaoxia, SUN Yugeng, CHEN Zengqiang, et al. Study of predicting network traffic using fuzzy neural networks [J]. Journal on Communications, 2005, 26(3): 136-140.
- [4] YAO Shuping, HU Changzhen, SUN mingqian. Prediction of web traffic based on wavelet and neural network [C]//Proceedings of IEEE WCICA 2006. Piscataway, NJ, USA: IEEE, 2006: 4026-4028.
- [5] 王俊松, 高志伟. 基于 RBF 神经网络的网络流量建模及预测[J]. 计算机工程与应用, 2008, 44(13): 6-11.  
WANG Junsong, GAO Zhiwei. Network traffic modeling and prediction based on RBF neural network [J]. Computer Engineering and Application, 2008, 44(13): 6-11.
- [6] MAO Guoqiang, LIU Huabing. Real time variable bit rate video traffic prediction [J]. International Journal of Communication Systems, 2007, 20(4): 491-505.
- [7] 陈剑, 闻英友, 赵大哲, 等. VBR 视频流量的小波包分解及其长时预测[J]. 通信学报, 2008, 29(6): 34-42.  
CHEN Jian, WEN Yingyou, ZHAO Dazhe, et al. Long-term prediction for VBR video traffic based on wavelet packet decomposition [J]. Journal on Communications, 2008, 29(6): 34-42.
- [8] YIN Hao, LIN Chuang, BERTON S, et al. Network traffic prediction based on a new time series model [J]. International Journal of Communication Systems, 2005, 18(8): 711-729.
- [9] SHU Yantai, JIN Zhigang, ZHANG Lianfang, et al. Traffic prediction using FARIMA models [C]//Proceedings of IEEE ICC'99. Piscataway, NJ, USA: IEEE, 1999: 891-895.
- [10] 吕军. 自相似网络流量的处理和分析 [J]. 清华大学学报(自然科学版), 2008, 48(7): 1186-1189.  
LÜ Jun. Analysis and processing of self-similar network traffic data [J]. Journal of Tsinghua University (Science and Technology), 2008, 48(7): 1186-1189.
- [11] KOLAROV A, ATAI A, HUI J. Application of Kalman filter in high-speed networks [C]//Proceedings of IEEE GLOBECOM'94. Piscataway, NJ, USA: IEEE, 1994: 624-628.
- [12] 李捷, 侯秀红, 韩志杰. 基于卡尔曼滤波和小波的网络流量预测算法研究 [J]. 电子与信息学报, 2007, 29(3): 725-728.  
LI Jie, HOU Xiuhong, HAN Zhijie. Application of Kalman filter and wavelet in traffic prediction [J]. Journal of Electronics and Information Technology, 2007, 29(3): 725-728.
- [13] SAGE A P, HUSA G W. Adaptive filtering with unknown prior statistics [C]//Proceedings of Joint Automatic Control Conference. Washington DC, USA: American Automatic Control Council, 1969: 760-769.

(编辑 刘杨)