

乘积季节模型在软件老化评估中的应用研究

李焱^{1,2,3}, 高强², 王勇², 刘欣然²

(1. 中国科学院计算技术研究所 北京 海淀区 100190; 2. 国家计算机网络应急技术处理协调中心 北京 朝阳区 100029;
3. 中国科学院大学 北京 海淀区 100049)

【摘要】在需要长期运行的系统中, 软件老化是一种常见的现象, 现有基于时序分析的软件老化评估方法, 大多基于简单的自回归或ARMA模型, 没有充分考虑软件老化关键指标的非平稳性、季节性等特征。该文提出一种基于乘积季节ARIMA模型的软件老化评估方法。并通过实验表明, 该方法能够较好地拟合季节性负载系统的软件老化趋势, 并能做出准确的预测以支撑软件再生。

关键词 乘积季节模型; 软件老化; 软件再生; 时序分析

中图分类号 TP311 文献标志码 A doi:10.3969/j.issn.1001-0548.2017.03.017

Software Aging Evaluation Method Using Multiplicative Seasonal ARIMA Model

LI Yan^{1,2,3}, GAO Qiang², WANG Yong², and LIU Xin-ran²

(1. Institute of Computing Technology, Chinese Academy of Sciences Haidian Beijing 100190;
2. National Computer Network Emergency Response Technical Coordination Center Chaoyang Beijing 100029;
3. University of Chinese Academy of Sciences Haidian Beijing 100049)

Abstract Software aging is a common phenomenon in a system that needs long-term operation. The existing analysis methods based on time series analysis mainly focus on autoregressive moving average (ARMA) models, not fully considered the seasonality or non-stationarity of the key indicators about software aging. This paper proposes a new software aging evaluation method based on seasonal autoregressive integrated moving average (ARIMA) model. The experimental results show that the method can well fit the software aging trend of seasonal load systems, and can achieve accurate prediction for supporting software rejuvenation.

Key words multiplicative seasonal model; software aging; software rejuvenation; time series analysis

已有研究表明, 当软件连续运行较长时间时, 其内部的一些错误条件的积累会导致软件性能衰退甚至停止运行, 这种现象称为软件老化^[1]。软件老化在需要长时间运行的系统中尤为常见, 如虚拟机和虚拟机管理器^[2]、操作系统^[3]、云平台系统软件^[4]、Java虚拟机^[5]、Web服务器^[6-7]、视频点播系统^[8]等。软件老化降低了系统的可靠性与可用性, 有时可能会给应用带来难以估量的损失。为了降低甚至消除软件老化带来的风险, 文献[1]提出了一种名为软件再生的方法, 及时评估系统老化状态并适时采取再生操作, 提升了系统可靠性。如何选择再生的时机是消除软件老化风险的关键, 再生太早则代价过高、效果不明显, 而再生太晚则起不到降低风险的作用。

确定再生时机的方法主要分为两种: 基于模型的方法和基于测量的方法^[6]。前者通过马尔科夫、随机Petri网等方法对系统建模, 计算出系统的最佳再生周期; 后者基于系统核心指标的观察数据进行软件老化评估, 预测关键资源耗尽或是降低到一定阈值的时间, 从而适时采取再生操作。基于测量的分析方法主要包括: 局部线性回归、决策树、支持向量机、时间序列分析等。与基于模型的方法相比, 基于测量的方法更侧重于反映真实观察数据的变化规律, 也更容易在工程领域被验证^[9]。可用内存数、内存交换区使用大小、CPU利用率等是已有基于测量方法的主要观察指标, 而上述指标的观察值和系统负载强相关。

收稿日期: 2015-11-25; 修回日期: 2016-02-24

基金项目: 国家973重点基础研究发展规划项目(2011CB302605); 国家科技支撑计划(2012BAH47B04)

作者简介: 李焱(1984-), 博士生, 主要从事分布式计算、云计算方面的研究。

值得关注的是,软件系统的老化趋势很大程度上受负载变化的影响^[6],负载越重、老化趋势会越快。而需要长时间运行的系统其负载通常表现出一定的周期性变化趋势,呈现出time-of-day或month-of-year效应^[10]。如基于Web服务器的ERP系统负载与员工上下班周期强相关,白天工作时间系统负载较重,员工下班后系统负载会逐步降低。但现有基于测量的方法未充分考虑系统负载的周期性变化特征。文献[9]主要针对目前软件老化分析中的单参数模型及未考虑变量间关联性和影响性的多参数模型的不足,提出了运用多元时间序列模型分析软件老化的方法,但未考虑系统负载特征与变量观测值之间的关系;文献[11]采用了响应时间、吞吐率、内存使用量等多个参数以反应软件老化趋势,并基于机器学习(MSP)算法预测软件的宕机时间,但同样对负载特征考虑不足。文献[6]通过观察内存交换区使用数、空闲物理内存数等指标变化情况,建立了季节性自回归(seasonal auto-regression)模型描述软件老化,同时给出了基于已使用内存交换区大小的预测结果,但AR模型本质上是一种针对平稳时间序列的分析方法,使得季节性自回归模型在考虑季节性负载的同时忽视了其变化的不平稳性。

因此,本文提出了一种基于乘积季节性自回归求和滑动平均模型(seasonal autoregressive integrated moving average, SARIMA)的软件老化评估方法。该方法充分考虑了系统负载给软件老化带来的影响,既涉及了季节性因素,又纳入了对动态趋势的评估,能够更为准确地预测关键指标(资源使用率)的变化趋势,以提升软件老化评估的效果。最后,基于一个实际线上系统的监测数据开展实验分析,说明了该方法的有效性。

1 乘积季节模型

时间序列分析^[12]是一种数据处理方法。该方法基于随机过程理论和数理统计学方法,研究随机数据序列所遵从的统计规律,以用于解决实际问题。时间序列分析在金融领域^[13]、机械制造^[14]、计算机领域^[6,9,15-18]等领域已有较为广泛的应用。文献[6,9]更是直接将时序分析方法用于软件老化的评估。

1.1 ARIMA模型

自回归求和滑动平均模型(ARIMA),是文献[13]于20世纪70年代初提出的一种著名的时间序列预测方法,所以又称为Box-Jenkins模型。该方法通过时

间序列的历史数据揭示数据随时间变化的规律,并将这种规律延伸到未来,从而具备一定的预测功能。该模型通常可表示为ARIMA(p,d,q),其中, p 为自回归项数, q 为移动平均项数, d 为使动态时间序列成为平稳序列所做的差分次数。常见的移动平均过程(MA)、自回归过程(AR)、自回归移动平均过程(ARMA)等都是ARIMA模型的特例。设 $X=(X_1,X_2,X_3,\dots,X_{n-1},X_n)$ 为一个时间序列,则ARIMA模型可表示为:

$$\begin{cases} \Phi(B)\nabla^d X_t = \Theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma^2, E(\varepsilon_t, \varepsilon_s) = 0, s \neq t \\ E(X_s, \varepsilon_t) = 0, \forall s < t \end{cases} \quad (1)$$

$$\nabla^1 X_t = (1-B)X_t = X_t - X_{t-1} \quad (2)$$

$$\Phi(B) = 1 - \varphi_1 B - L - \varphi_p B^p \quad (3)$$

$$\Theta(B) = 1 - \theta_1 B - L - \theta_q B^q \quad (4)$$

式中, B 为后移算子; ∇^d 表示向后差分算子; t 为时间标志; $\{\varepsilon_t\}$ 表示白噪声序列; d 为差分次数。若取一阶差分,即 $\nabla^1 = 1-B$,则有:

$$\nabla^1 X_t = (1-B)X_t = X_t - X_{t-1} \quad (5)$$

式(3)中 $\Phi(B)$ 阶数为 p ,式(4)中 $\Theta(B)$ 阶数为 q ,也即模型ARIMA(p,d,q)。

1.2 乘积季节模型

对于非平稳的时间序列(如我国GDP值组成的时间序列,一直呈上升趋势),普通的ARIMA模型通过 d 阶差分已能很好的模拟其变化规律。

但是,在某些时间序列中,存在明显的周期性变化。这种周期是由于季节性变化(包括天、周度等变化)或其他一些固有的因素引起的,这类序列称为季节性序列。如一个地区的气温值序列(每隔一小时取一个观测值)中除了含有以天为周期的变化,还含有以年为周期的变化。在计算机领域,季节性序列也比较常见,如铁路购票站点访问量序列、网上购物站点负载等。处理季节性时间序列只用以上介绍的方法是不够的,描述这类序列的模型之一是乘积季节ARIMA模型。该模型不仅考虑了不同周期中相同周期点之间的相关性,还考虑了相同周期内不同周期点之间的关系,通常比较符合时间序列观测数据的实际情况。乘积季节ARIMA模型可表示为:

ARIMA(p,d,q)(k,D,m) $_S$,即:

$$\Phi(B)U(B^S)\nabla^d \nabla_S^D X_t = \Theta(B)V(B^S)\varepsilon_t \quad (6)$$

$$U(B^S) = 1 - \Gamma_1 B^S - \Gamma_2 B^{2S} - \dots - \Gamma_k B^{kS} \quad (7)$$

$$V(B^S) = 1 - H_1 B^S - H_2 B^{2S} - \dots - H_m B^{mS} \quad (8)$$

$$\nabla_S^D = (1 - B^S)^D \tag{9}$$

式中, S 为单位周期内采样点的个数; k 为季节性自回归阶数; D 为季节性差分阶数; m 为季节性移动平均阶数。式(6)中 ∇^d 、 $\Phi(B)$ 、 $\Theta(B)$ 的表达式分别和式(2)、式(3)、式(4)一致。 $\Phi(B)$ 和 $\Theta(B)$ 消除同一周期的不同周期点之间的相关性; $U(B^S)$ 和 $V(B^S)$ 消除不同周期的同一周期点之间的相关性。

2 基于乘积季节模型的老化评估

基于测量的软件老化评估方法首先需选取反应系统老化状态的关键指标, 如CPU使用、内存使用、内存交换区使用等相关指标。然后, 基于指标的当前及历史测量值, 利用某种模型或算法来预计指标未来值的变化情况, 从而得出最佳的再生时间。本文提出的基于乘积季节模型的老化评估方法如图1所示。主要有以下3个步骤:

1) 确定老化指标。分析系统历史监测数据, 得出反应系统老化的关键指标, 常见的如CPU利用率、内存使用、交换区使用等。

2) 建立模型。按照老化指标在历史监测数据中提取相应的训练数据, 并建立具体老化评估模型。模型建立主要分为3个阶段: 模型识别、模型参数估计和模型诊断。模型识别需要在训练数据上观察自相关数(ACF)、偏自相关数(PACF)等指标情况, 并依据AIC或BIC准则判断模型的具体类型^[13]; AIC(akaikie information criterion)是赤池信息准则, 它重点兼顾了模型的简洁性和精确性, 鼓励数据拟合的优良性但尽量避免出现过度拟合现象。BIC(Bayesian information criterion)是贝叶斯信息准则, 它根据贝叶斯信息量衡量模型的优良; 模型参数估计确定模型的具体参数; 模型诊断检验模型在测试数据上的拟合效果, 若拟合结果较差, 则需要重新进行模型识别、参数估计等操作。目前, Eviews软件、R语言工具^[19]已提供了较为方便的手段完成上述步骤。

3) 老化评估。模型预测即完成软件老化评估的功能。提取实时监测数据的老化特征后作为模型的输入, 得出的输出即为老化特征的预测结果。实际可结合需求进行多步预测, 以留出充足的时间执行再生操作。但随着预测步数的增长, 准确率会降低。

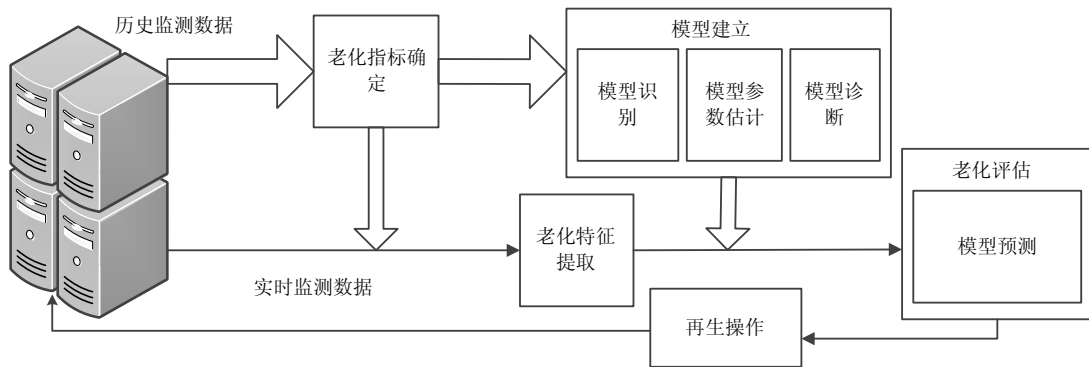


图1 老化评估方法

3 实验与评估效果分析

3.1 数据集

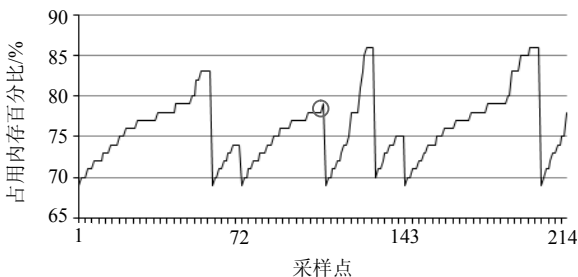


图2 某系统内存占用情况

为了说明本文老化评估方法的有效性, 收集了某线上系统连续3天的实际运行监测数据。主要包括CPU占用率、内存占用率等指标。为避免影响实际

系统的运行, 降低了监测数据的采集频率。最终采集频率为20 min/次, 数据规模为216条。

在反映软件老化状况的众多指标中, 内存使用是其中的一项关键指标^[3,6,8]。因此, 本文把内存作为软件老化的特征展开分析。数据集中关于内存的统计如图2所示。可以看出, 内存消耗具备一定的季节性特征。每个周期开始后, 内存占用率会持续上升, 当上升到一定值后会骤降(对应时间点定时重置系统所致), 而后又会继续上升。其中, 第2个周期内圆圈标识点内存使用骤降是因为实际故障处置中的重启所致。图2的数据表明, 随着系统运行时间的增加, 内存使用率指标上升趋势逐步加快, 系统老化现象也愈发明显。同时, 软件再生(重启来实现)及时降低内存消耗, 避免因内存不足导致软件崩溃。

3.2 效果分析

本文借助R语言工具(3.1.2版本)^[19]完成乘积季节模型识别、参数估计等操作,并通过模型诊断、关键指标预测等说明了本文方法的有效性。

3.2.1 模型识别与参数估计

图2中的统计数据表现出了很强的季节性特征,因此,首先能够确定模型为一种季节性模型。而后,对同一周期内数据特征进行分析,主要是计算其ACF(自相关)值和PACF(偏自相关)值。

实际计算得知,内存数据具有较强的非平稳特征,无明显的截尾或拖尾现象。截尾是统计数据在某个点之后都突然为零的现象,拖尾是统计数据在某个点之后统计值逐步降低并很快接近于零的现象。需要通过差分操作,将统计数据变成平稳的。对一般的非平稳时间序列,一次差分就能使之变成平稳的时间序列。即先确定了模型为ARIMA(p, I, q)(k, l, m)₇₂,并通过观察差分后平稳序列的自相关和偏自相关图,得出可能性最大的模型为:ARIMA(0,1,0)(0,1,0)₇₂, ARIMA(0,1,1)(0,1,0)₇₂, ARIMA(1,1,0)(0,1,0)₇₂, ARIMA(1,1,1)(0,1,0)₇₂, ARIMA(0,1,0)(0,1,1)₇₂, ARIMA(0,1,1)(0,1,1)₇₂, ARIMA(1,1,0)(0,1,1)₇₂, ARIMA(1,1,1)(0,1,1)₇₂。

最终模型的选定主要依据AIC和BIC最小准则。

上述模型的对应指标统计值如表1所示。

表1 各模型参数比较

模型	AIC值	BIC值
ARIMA(0,1,0)(0,1,0) ₇₂	255.68	259.94
ARIMA(0,1,1)(0,1,0) ₇₂	255.29	259.82
ARIMA(1,1,0)(0,1,0) ₇₂	255.36	259.88
ARIMA(1,1,1)(0,1,0) ₇₂	256.57	263.36
ARIMA(0,1,0)(0,1,1) ₇₂	256.68	261.2
ARIMA(0,1,1)(0,1,1) ₇₂	257.29	264.08
ARIMA(1,1,0)(0,1,1) ₇₂	257.06	264.83
ARIMA(1,1,1)(0,1,1) ₇₂	258.54	267.59

基于表1的结果,结合AIC和BIC准则,最终选取的模型为ARIMA(0,1,1)(0,1,0)₇₂。

3.2.2 模型诊断

对上述模型进行拟合,得到在前两个周期里真实值和拟合值的时序图,如图3所示。

图3中,实线为真实值,虚线为模型拟合值。该图显示出拟合值能较好拟合真实统计数据。另外,也可对模型的残差序列进行白噪声检验以判断模型的真实效果。一般地,若残差序列是白噪声序列,则认为该模型适合模拟当前时间序列,若残差序列

不是白噪声序列,则认为还存在有用信息未提取出来,模型还需进一步改进。用R语言中的Box.test()函数完成残差序列的随机性检验,最终的 p 值为0.875,可以认为残差序列为白噪声序列。同时,还可以检验残差序列的自相关数随着滞后阶数的增大是否逐渐趋近于零。检验结果如图4所示,残差序列相关系数均落入了随机区间,且基本上都小于0.2。因此,ARIMA(0,1,1)(0,1,0)₇₂模型较好地拟合了真实数据。

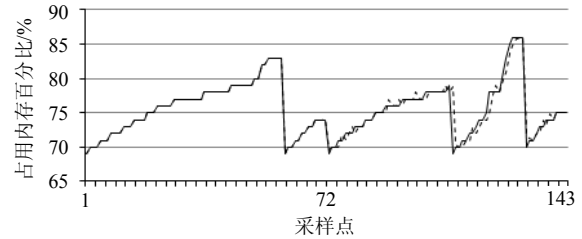


图3 真实值和模型拟合值对比

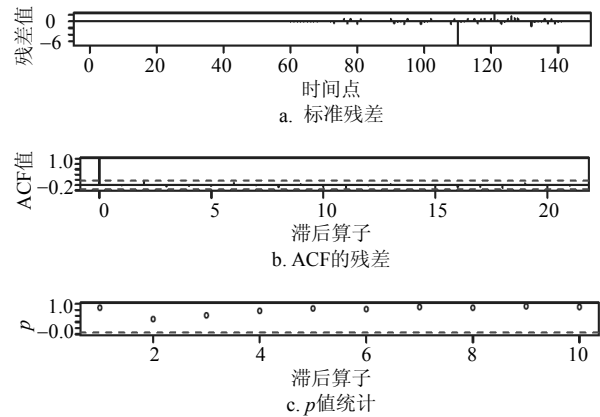


图4 模型检验结果

3.2.3 效果分析

软件老化评估是根据关键指标的历史时间序列,运用乘积季节ARIMA模型对关键指标未来值进行推测。若该值在未来某时间点上的推测结果超过一定的阈值,则认为系统将会在该点发生异常,系统管理员可提前采取备份、软件再生等操作,以降低老化给业务运行带来的潜在影响。

本节利用上述数据集进行预测分析。把前两个周期的数据当作是历史数据,第3个周期内的数据当作是测试数据,以比对模型的预测效果。预测阶数为12,每个数据点采集频率为20 min,因此提前预测的时长为240 min(4 h)。第一点预测效果如图5所示(实线为测试集中的实际值,虚线为对应点上的预测值)。

从图5可以看出,预测曲线的走势与实际数据曲线的走势基本一致。另外,本文还计算了预测结果

的MAPE值(the mean absolute percentage error), 即平均绝对误差百分比来对模型进行评价。上述12阶预测的MAPE值为0.77%, 即预测的相对误差控制在1%以内, 能够很好地满足再生策略的生产需求。

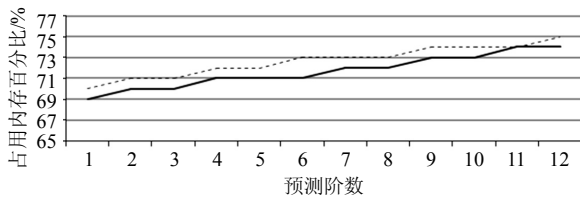


图5 预测值与实际值比较图

4 结束语

如何评估软件系统的老化状况以适时采取再生操作是软件老化研究的核心问题之一。本文提出了一种基于乘积季节ARIMA模型的软件老化评估方法, 该方法基于时序分析理论, 在考虑软件老化关键指标变化的非平稳性的同时, 利用了指标变化的季节性趋势, 能够准确地拟合系统实际负载情况。并基于某个线上系统实际监测数据集, 通过实验分析, 说明了该方法的有效性。未来的工作将集中于软件老化特征的提取及不同特征相关性的分析上, 以更为全面地评估软件老化状况, 从而生成更为科学的软件再生策略。

参 考 文 献

- [1] HUANG Y, KINTALA C, KOLETTIS N, et al. Software rejuvenation: Analysis, module and applications[C]// Twenty-Fifth International Symposium on Fault-Tolerant Computing, 1995, FTCS-25. [S.l.]: IEEE, 1995: 381-390.
- [2] MACHIDA F, XIANG J, TADANO K, et al. Combined server rejuvenation in a virtualized data center[C]//2012 9th International Conference on Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC). [S.l.]: IEEE, 2012: 486-493.
- [3] DOMENICO C, ROBERTO N, ROBERTO P, et al. Software aging analysis of the linux operating system[C]// In Proc of Int'l Symp on Software Reliability Engineering (ISSRE2010). [S.l.]: IEEE, 2010: 71-80.
- [4] ARAUJO J, MATOS R, MACIEL P, et al. Experimental evaluation of software aging effects on the eucalyptus cloud computing infrastructure[C]//Proceedings of the Middleware 2011 Industry Track Workshop. [S.l.]: ACM, 2011: 4.
- [5] COTRONEO D, ORLANDO S, RUSSO S. Characterizing aging phenomena of the java virtual machine[C]//26th IEEE International Symposium on Reliable Distributed Systems, 2007, SRDS 2007. [S.l.]: IEEE, 2007: 127-136.
- [6] GROTTKE M, LI L, VAIDYANATHAN K, et al. Analysis of software aging in a web server[J]. IEEE Transactions on Reliability, 2006, 55(3): 411-420.
- [7] LI L, VAIDYANATHAN K, TRIVEDI K S. An approach for estimation of software aging in a web server[C]//Empirical Software Engineering, International Symposium. [S.l.]: IEEE, 2002: 91-100.
- [8] 杜小智, 齐勇, 鲁慧民, 等. 视频点播系统的软件老化估计和预测[J]. 计算机研究与发展, 2012, 48(11): 2139-2146.
DU Xiao-zhi, QI Yong, LU Hui-min, et al. Software aging pattern analysis of the video on demand system[J]. Journal of Computer Research and Development, 2011, 48(11): 2139-2146.
- [9] 郑鹏飞, 齐勇, 陈鹏飞. 软件老化的多元时间序列分析方法[J]. 计算机科学与探索, 2012, 6(2): 125-133.
ZHENG Peng-fei, QI Yong, CHEN Peng-fei. Multivariate time series analysis of software aging[J]. Journal of Frontiers of Computer Science and Technology, 2012, 6(2): 125-133.
- [10] WILLIAMS A, ARLITT M, WILLIAMSON C, et al. Web workload characterization: Ten years later[M]//Web Content Delivery. [S.l.]: Springer US, 2005: 3-21.
- [11] ALONSO J, TORRES J, BERRAL J, et al. Adaptive on-line software aging prediction based on machine learning[C]//Proceedings of the IEEE/IFIP International Conference on Dependable Systems and Networks. Washington, DC, USA: IEEE Computer Society, 2010: 507-516.
- [12] CHATFIELD C. The analysis of time series: an introduction[M]. [S.l.]: CRC, 2013.
- [13] 巩师恩, 范从来. 二元劳动力结构与通货膨胀动态形成机制——基于新凯恩斯菲利普斯曲线框架[J]. 财经研究, 2013, 3: 75-86.
GONG Shi-en, FAN Cong-lai. Dual labor structure and dynamic formation mechanism of inflation: Based on the framework of new keynesian phillips curve[J]. Journal of Finance and Economics, 2013, 3: 75-86.
- [14] CHEN K Y, WANG C H. A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan[J]. Expert Systems with Applications, 2007, 32(1): 254-264.
- [15] TSENG F M, TZENG G H. A fuzzy seasonal ARIMA model for forecasting[J]. Fuzzy Sets and Systems, 2002, 126(3): 367-376.
- [16] 袁小坊, 陈楠楠, 王东, 等. 城域网应用层流量预测模型[J]. 计算机研究与发展, 2009, 46(3): 434-442.
YUAN Xiao-fang, CHEN Nan-nan, WANG Dong, et al. Traffic prediction models of traffics at application layer in metro area network[J]. Journal of Computer Research and Development, 2009, 46(3): 434-442.
- [17] 黄建国, 罗航, 王厚军, 等. 运用GA-BP神经网络研究时间序列的预测[J]. 电子科技大学学报, 2009, 38(5): 687-692.
HUANG Jian-guo, LUO Hang, WANG Hou-jun, et al. Prediction of time sequence based on GA-BP neural net[J]. Journal of University of Electronic Science and Technology of China, 2009, 38(5): 687-692.

(下转第611页)

- 2005: 924-931.
- [7] HOSNI A, BLEYER M, GELAUTZ M, et al. Local stereo matching using geodesic support weights[C]//IEEE International Conference on Image Processing. [S.l.]: IEEE, 2009: 2069-2072.
- [8] YANG Q. A non-local cost aggregation method for stereo matching[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Computer Society, 2012: 1402-1409.
- [9] MEI X, SUN X, DONG W, et al. Segment-tree based cost aggregation for stereo matching[C]//Conference on Computer Vision and Pattern Recognition (CVPR). [S.l.]: IEEE, 2013: 313-320.
- [10] COMANICIU D, MEER P. Mean shift: a robust approach toward feature space analysis[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(5): 603-619.
- [11] RHEMANN C, HOSNI A, BLEYER M, et al. Fast cost-volume filtering for visual correspondence and beyond[C]//Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Computer Society, 2011: 3017-3024.
- [12] SCJARSTEIND, SZELISKI R. Middlebury stereo evaluation[EB/OL]. [2015-12-20]. <http://vision.middlebury.edu/stereo/eval/>.

编辑 漆蓉

(上接第587页)

- [18] 吴少智, 吴跃, 徐鹏, 等. 支持向量回归的颅内压时间系列无损估计方法[J]. 电子科技大学学报, 2011, 40(6): 956-960.
WU Shao-zhi, WU Yue, XU Peng, et al. Support vector regression based time series mining approach for non-invasive ICP assessment[J]. Journal of University of Electronic Science and Technology of China, 2011, 40(6): 956-960.
- [19] R-Project. Rcasting functions for time series and linear models[EB/OL]. [2015-11-01]. <http://cran.r-project.org/web/packages/forecast/index.html>.

编辑 漆蓉